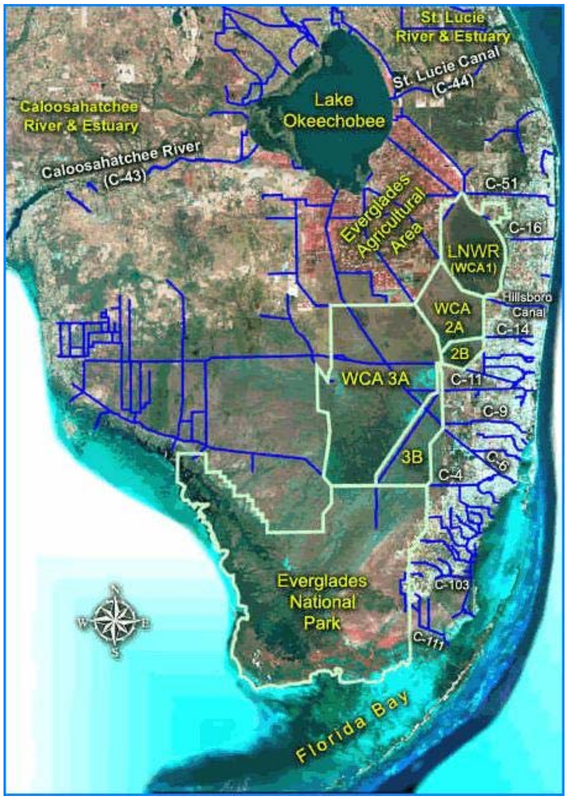
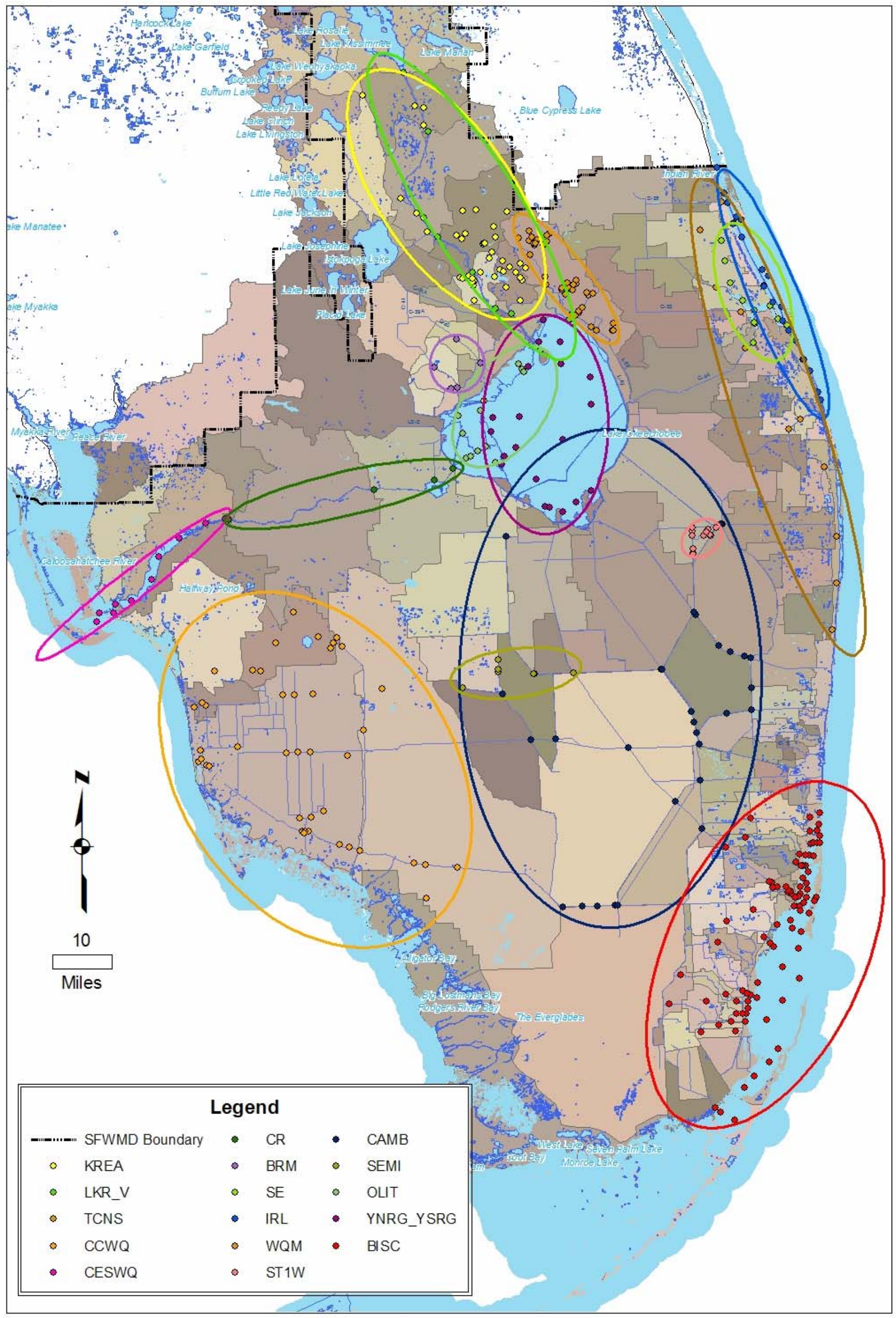
Surface Water Quality Monitoring Network Optimization

Comprehensive Report to the South Florida Water Management District





Final

February 2006

Battelle

SURFACE WATER QUALITY MONITORING NETWORK OPTIMIZATION COMPREHENSIVE REPORT

Work Order Number C-15968-WO04-11

South Florida Water Management District
Environmental Resource Assessment Department
1000 N.E. 40th Ave.
Okeechobee, FL 32963

Final

Carlton D. Hunt[1], Ph.D.
Jennifer Field[1], M.S.
Steve Rust[2], Ph.D.
Patricia Burke[3], M.S.

February 24, 2006

[1]Battelle
397 Washington Street
Duxbury, MA 02332
(617) 934-0571

[2]Battelle
505 King Avenue
Columbus OH 43201

[3]South Florida Water Management District
Dept. of Monitoring and Assessment
8894 Belvedere Rd.
West Palm Beach, Fl 33411

Battelle
The Business *of* Innovation

TABLE OF CONTENTS

ATTACHMENTS

LIST OF TABLES

ACKNOWLEDGEMENTS

Many individuals from multiple organizations contributed to the statistical approach and project-specific optimizations. Mr. David Wade and Mr. Mike Wessel from Janicki Environmental worked with Dr. Steve Rust and Battelle statisticians Fred Todt and Matt Sanders to develop the approach used for the optimizations. Mr. Wade and Mr. Wessel were also responsible for optimization of projects SE, IRL, WQM, CR, CESWQ, CCWQ as well as BRM and STA1W with assistance from Probas Adack from Barnes Ferland and Associates. Along with the statistical protocol, Dr. Rust was responsible for performing the optimizations for LKR, V, TCNS and KREA projects. Ms. Jennifer Field focused on the OLIT and Y project optimizations and Dr. Carlton Hunt conducted the Optimization for project BISC. Mr. Skip Newton conducted optimizations for the SEMI and CAMB projects.

District staff and members of the Districts' Optimization Technical Team provided valuable input for the optimization projects. The list of projects and contributing staff are.

CR: Dr. Peter Doering, Mr. Dan Crean, Ms. Carole Maddox
CESWQ: Dr. Peter Doering, Mr. Dan Crean, Ms. Carole Maddox
CCWQ: Mr. Ananta Nath, Mr. Mike Duever, Mr. Tim Howard, Ms. Carole Maddox
BRM: Ms. Patricia Burke, Mr. Pat Davis
SE: Mr. Dan Crean, Dr. Nenad Iricanin, Ms. Monique Laham-Pass
IRL: Mr. Dan Crean, Dr. Nenad Iricanin, Ms. Monique Laham-Pass
WQM: Mr. Dan Crean, Dr. Nenad Iricanin, Dr. Yongshan Wan, Ms. Monique Laham-Pass
LKR: Ms. Steffany Gornack, Mr. Brad Jones, Mr. Pat Davis
V: Ms. Steffany Gornack, Mr. Brad Jones, Mr. Pat Davis
KREA: Mr. Gary Ritter, Ms. Steffany Gornack, Dr. Joyce Zhang, Ms. Brad Jones, Mr. Pat Davis
TCNS: Mr. Gary Ritter, Ms. Steffany Gornack, Dr. Joyce Zhang, Ms. Brad Jones, Mr. Pat Davis
OLIT: Dr. Tom James, Mr. Bruce Sharfstein, Mr. Chuck Hanlon, Mr. Pat Davis
Y: Dr. Tom James, Mr. Bruce Sharfstein, Mr. Chuck Hanlon, Mr. Pat Davis
ST1W: Dr. Jana Newman, Ms. Kristin Larson
CAMB: Ms. Linda Crean, Ms. Barb Powell, Mr. Stuart Van Horn, Mr. Larry Grosser, DR. Cheol Mo, Mr. Patrick Martin, Ms. Patricia Burke
SEMI: Ms. Barb Powell, Mr. Stuart Van Horn, Mr. Larry Grosser
BISC: Dr. Dave Rudnick, Mr. Steve Kelly, Ms. Robin Bennet, Ms. Teresa Coley, Dr. Nenad Iricanin, Ms. Danielle Tharin, Dr. Stephen Blair (DERM) Dr. Susan Markley (DERM) and Dr. Joe Boyer (FIU).

EXECUTIVE SUMMARY

The South Florida Water Management District (District) is continuously challenged with providing the resources needed to accommodate substantial and diverse water quality data needs. With over 1500 active monitoring sites, the District's surface water quality monitoring network spans a wide variety of ecosystems over a large geographic area. The network consists of several individual monitoring projects (groupings of monitoring station/sites) driven by a diverse set of mandates (*i.e.,* laws, permits, agreements, etc.) and objectives. This monitoring must be accomplished under the constraint of priority initiatives being supported by limited financial resources and manpower. To ensure cost effective monitoring, improve service, and position the District to accommodate future monitoring requirements, the Environmental Resource Assessment Department conducted a detailed optimization of its non-permit driven water quality monitoring program. The findings of the optimization, as well as recommendations for evaluating future monitoring initiatives are reported in this document.

The optimization effort was modeled after the U.S. Environmental Protection Agency's Data Quality Objective (DQO) process and used robust statistical analyses to evaluate and individually optimize 23 of the District's current Type II and III mandated surface water quality monitoring projects. Each project optimization focused on the how stations were selected, what drives the particular parameter set at each station and the justification or logistics relating to sampling frequency. This was accomplished by clearly identifying how the project data were used (i.e., data end use), which mandates affect which station, the relevance of each monitoring project's goals and objectives to the sampling stations and how these factors relate to the District's mission. Five of the original projects were identified that either were too early in their life cycle to be optimized, were research projects, or were simply not amenable to statistical optimization. Several other specific projects were combined for the optimization activity. This resulted in eighteen (18) projects that were analyzed using statistical methods and/or the DQO process. Data gathered by these projects from 1992 through 2003 were accessed from DBHydro and reviewed to develop District approved project specific databases for the statistical analysis. The projects were optimized using up to five key water quality parameters monitored under the project.

Recommendations for a standard approach to evaluating current water quality monitoring projects and considerations for future monitoring projects were also developed. The District currently develops and maintains water quality monitoring efforts within the frame work of the DQO process and it is recommended that the components of this vital practice continue to be applied by the District. Components critical to the success of monitoring initiatives are the development of clear statements of the project monitoring goals and objectives, a complete description of all data uses, and a thorough awareness of the management and policy decisions the data will support. Moreover, addressing seasonal trends and autocorrelation in monitoring project data are crucial to ensure statistical results are not overstating the power of the monitoring to detect trends.

In keeping with the recommendations, the data uses for each project were compiled and reviewed. For most projects, the key uses of the data coalesced around the ability to detect trends and changes from those trends. Therefore, a statistical program that relies on methods from the Mann-Kendall trend analysis developed by Reckhow *et al.,* 1993 was developed in SAS. The SAS code and step by step instructions for this procedure was supplied to the District for use it for future optimizations and for determining monitoring designs for new programs.

Specific projects were optimized on the potential for improving the following: evaluation to identify any redundant stations; evaluation of current sampling frequency against alternative frequencies; and evaluation of specific parameters in the project monitoring program. Optimization recommendations for each specific project ranged from modification of sampling frequency to revisions in the stations sampled (remove, relocate, add), to changes in the parameters measured. The period of record for several projects (BRM, CCWQ, CESWQ, and IRL) was too short to conduct adequate statistical optimizations. Continued monitoring and application of the optimization process are recommended for these projects. Moreover, continued discussion with other stakeholders and partners is encouraged to ensure data generated by the Districts program are value added and useful for management of water in south Florida.

The individual project optimization findings and recommendations coalesced around several general themes relevant to the District's water quality optimization initiative. These include 1) clearly defining the data end uses so that the monitoring program can be designed to collect the appropriate information; 2) expanding the optimization for several projects from concentration based data to loading which incorporates flow data (i.e., optimize on loading rather than concentration); 3) maximizing the use of autosamplers where loading data is a key end use (reducing effort on back up grab samples; 4) removal of several parameters that District staff indicated are no longer used; and 5) better describe the amount of change and time period for change detection to improve temporal optimization. Moreover, the statistical results suggest that at least ten years of data are required to detect trends, indicating commitments to monitoring need to be sustained.

From the perspective of sample collection frequency, high temporal variability and fixed seasonal effects along with the high degree of autocorrelation limits the ability to obtain truly independent samples. Unless these data issues are considered correctly, they provide overly optimistic estimates of the ability to detect trends. These issues require additional considerations within each project and the District in general.

The magnitude of cost savings realized by implementation of the recommendations contained in this report is difficult to quantify. The costs associated with the monitoring programs can be broken into four categories: field costs, laboratory costs, quality assurance cost, and data loading/maintenance costs. Many of the changes recommended incrementally modify labor efforts, thus must be evaluated within context of the entire water quality measurement program and balanced against the value gained. The District has instituted a successful cost saving practice for monitoring programs by forming cost sharing partnerships with local, state, federal and tribal interests. It is recommended that the District continue to investigate further opportunities to share the financial and resource needs of monitoring programs and partner with agencies that have similar needs for data and information.

1. INTRODUCTION

The South Florida Water Management District's (District) Environmental Resource Assessment (ERA) Department is currently responsible for ensuring that water quality data collected by the District complies with regulatory requirements and are defensible and of acceptable quality. The District is continuously challenged with providing the resources needed to accommodate substantial and diverse water quality data needs. With over 1500 active monitoring sites, the District's surface water quality monitoring network spans a wide variety of ecosystems over a large geographic area. The network consists of several individual monitoring projects (groupings of monitoring sites) driven by a diverse set of mandates (*i.e.,* laws, permits, agreements, etc.) and objectives. This monitoring must be accomplished under the constraint of priority initiatives being supported by limited financial resources and manpower. The costs, as well as the level of monitoring commitments (*i.e.,* permit specific requirements) undertaken by the District are expected to increase substantially in the future.

While the District has been able to internally review and optimize the majority of its monitoring sites, its ability to statistically optimize the water quality monitoring network frequently and on a project specific and wide scale basis became problematic due to commitments of staff to other high priority projects. The ERA Department was audited in 2004 as a partial response to the District's management goals of examining business processes. A recommendation of the audit was to contract a comprehensive optimization plan for its overall water quality monitoring program. Therefore, this initiative was established to optimize the bulk of the District's surface water quality monitoring program. This optimization effort is intended to incorporate statistically based optimization methods and also give equal weight to the need and relevance of each monitoring effort, as it relates to the District's mission and priority projects. The commitment to support a comprehensive optimization study is based on commitment to scientifically defensible recommendations that can be considered and implemented in a single step, thus allowing any cost savings and staffing availabilities to be visible and accessible over an accelerated time scale.

The District's current water quality monitoring projects generally fall into one of the following categories:

- **Type I Mandate***:* Work is required by state or federal statute or permit which is very specific and does not allow for much District discretion in implementation. *Examples*: NPDS Permit for STA-1W, Non-ECP Permit, Everglades Forever Act Permit, Operation of the C&SF System, etc.
- **Type II Mandate**: Work is required by statute, permit, regulation or agreement and allows for some discretion. *Examples*: Minimum Flows and Levels, Water Supply Plans, Everglades Consolidated Report, Regulatory Programs, CERP Projects, etc.
- **Type III Mandate**: Project is not legally mandated, but has been authorized or requested by the Governing Board, Executive Office, or another agency. This type of mandate may include ongoing research that is used to establish criteria and set parameters to obtain future funding to address relevant concerns. The District has complete discretion in implementation. *Examples*: SWIM Projects, most local cooperative agreements, expert assistance program, and research type monitoring that supports specific projects and District directives.

The District's Type I monitoring network was reviewed under a separate initiative. The findings from that study can not be changed as a result of the work conducted under this project. However, the rest of the District's surface water quality monitoring network must be evaluated with an understanding of all components. To optimize the Type II and III projects, there must also be a clear understanding and incorporation of what is being done at the Type I level. While system wide optimization is desired, this project only focuses on a set of specific projects and monitoring sites which have been identified by the district for optimization. These projects may incorporate Type I sites which will be considered along with the others sites during the optimization.

1.1 Project Goals and Objectives

The goal of this project was to use scientifically defensible methods and robust statistical analyses to evaluate and optimize 25 of the District's current Type II/III level water quality monitoring projects and to identify any associated costs or resource savings and benefits. Meeting this goal required a clear understanding of the reasons why each individual station within a project was monitored, which mandates affect which station, how the data from each station are used, the relevance of each monitoring project's goals and objectives to the sampling stations and water quality monitoring project, and how these factors relate to the District's mission. Early in the optimization process five projects were identified that either were too early in their life cycle to be optimized or that supported biological studies necessary to understand and predict the impacts of water management decisions. Projects that fell into these categories with their study focus are:

- Tree Island Monitoring (TREE) Project: Ground water and hydrology studies in support of predictions on tree island response to CERP activities;
- Florida Bay Monitoring Network (FLAB) Project: *In situ* water quality measurements only and these are used to support SAV monitoring and recovery in Florida Bay,
- Test Cells (TSTC) Project: This project was a research study within the Stormwater Treatment Area 1 West (STIW) project;
- Big Cypress Basin Water Quality (BCWQ) Project: The project was not continued due to lack of funding; and
- Lake Okeechobee Critical Projects (LOCP): Only one station is sampled therefore, the project is not amenable to statistical optimization.

Other project were defined by the sampling logistics required to complete sampling (ENRR and ENRU; YNRG and YSRG). These were combined as ST1W and Y, respectfully, for optimization. These actions reduced the number of projects receiving detailed analysis to eighteen (Table 1).

Table 1. Project Optimization List and Regional Assignment

Region	Project Acronym	Project Title
North of Lake Okeechobee	KREA	Kissimmee River Eutrophication Abatement Project
	LKR	Lower Kissimmee River
	TCNS	Taylor Creek Nubbin Slough
	V	Kissimmee River Structures
West of Lake Okeechobee	BRM	Brighton Reservation Monitoring
	CCWQ	Collier County Water Quality
	CESWQ	Caloosahatchee Estuary Water Quality
	CR	Caloosahatchee River Project
East of Lake Okeechobee	IRL	Indian River Lagoon
	SE	St. Lucie Estuary
	WQM	Lake Worth and West Palm Beach Monitoring Network
Lake Okeechobee	OLIT	Lake Okeechobee Littoral Zone
	Y	Lake Okeechobee - In Lake North and South
Everglades Agricultural Area	CAMB	Conservation Areas Inflows and Outflows
	RAIN	Rain and Atmospheric Deposition
	SEMI	Seminole Reservation
	ST1W	Stormwater Treatment Area 1 West
Southeast Coast	BISC	Biscayne Bay (DERM and FIU Monitoring programs)

1.2 Project Activities and Technical Approach

This optimization project was conducted under the seven tasks listed below.

- Task 1 – Kick-off meeting
- Task 2 – Questionnaire
- Task 3 – Literature Search
- Task 4 – Work Plan
- Task 5 – Progress Reporting
- Task 6 – Comprehensive Report
- Task 7 – Recommendations for developing a District Monitoring Evaluation Tool

The Work Plan developed under Task 4 detailed the approach that was used to optimize each of the projects listed in Table 1. Progress reporting under Task 5 consisted of six meetings with District staff to review progress, evaluate findings, and modify the technical direction of the project. These activities are summarized in progress reports and meeting summaries which are available under separate cover from the District. This report constitutes Task 6 of the work plan. A SAS based temporal optimization tool developed to support the optimization was delivered under separate cover and addresses Task 7. The description of the SAS routine is included as Attachment 1 of this comprehensive report and includes summary recommendations for a District Monitoring Evaluation Tool.

The remainder of this report is organized as follows: Section 2 describes the approach that was used to conduct the project optimizations. Section 3 briefly summarizes each of the monitoring projects that were optimized during this effort. Section 4 summarizes the optimization recommendations, benefits and partnership opportunities identified during the optimization. Section 5 summarizes comments on previous District optimization efforts reviewed as part of this project. Section 6 conveys the recommendations for a District Monitoring Evaluation Tool and summarizes the statistical SAS tool developed for trend optimization.

Individual Project Summaries containing the results of the statistical optimizations were submitted under separate cover. Final decisions regarding implementation of the recommendations are the responsibility of the District.

2. TECHNICAL APPROACH

The optimization process incorporated elements of the U.S. Environmental Protection Agency (EPA) Data Quality Objectives (DQO) Process as described in U.S. EPA (2000). The seven steps of the DQO process are:

1. **Step 1: State the problem**. Identify all legislative mandates or other monitoring goals and objectives that motivate the water quality monitoring being conducted.
2. **Step 2: Identify the decision**. Identify how currently collected monitoring information is used by the District to address the mandate or monitoring objective. (i.e., end uses of the data and reports that use the information from the monitoring program, models that need the monitoring data etc.)
3. **Step 3: Identify inputs**. Identify which parameters need to be measured.
4. **Step 4: Define boundaries**. Identify both spatial and temporal circumstances that must be considered to adequately address the decisions (i.e., define the geographical boundaries associated with the mandate or monitoring objective, identify the stations that are currently monitored within those geographical boundaries, identify the frequency of sampling).

5. **Step 5: Develop a decision rule**. Specify the statistical analysis procedures employed to support the District's use of the monitoring data as well as the manner in which analysis results are used to make District decisions.
6. **Step 6: Specify limits on decision errors**. Define acceptable levels of decision error or acceptable levels of uncertainty in statistical analysis results.
7. **Step 7: Optimize the design**. Perform statistical analysis on the current and alternative designs to identify which design is the most appropriate and can answer the question in the most cost effective manner.

These steps were consolidated into an Optimization Project Work Plan (Battelle 2005) which included a literature search and review of published documents on optimization. Steps 1-4 of the DQO process were used to develop a questionnaire to gather clarifying information on how data from each monitoring project supports the District decision-making processes. Information on the goals and objectives, mandates, how and where the data are used (reported), plus current and future District goals and objectives the data might support was also solicited. Other information sought was related to the specific parameters, sampling types, and general monitoring design issues, financial costs, benefits, and partnerships plus potential optimizations from the staff's perspective. This type of information was to be incorporated into project-specific summaries for each project undergoing optimization. Preliminary project summaries were prepared by examining existing District monitoring plans and compiling summary tables showing which stations were sampled as part of the project, which parameters were measured at those stations and with what frequency. The project summaries also communicated general information that was taken from existing documentation on the project including the project's start date, managers responsible for the project, spatial extent, and the purpose, goals, objectives, and mandates that supported the project.

After District staff reviewed the questionnaire, it was determined that a more effective approach would be to conduct one-on-one interviews with the District's project teams. This was found to be highly effective in helping to reconcile the information in the project summaries. Based on the information obtained from the interviews, the project summaries were updated to include more detailed information of specific goals and objectives, as well as information on how the data from each monitoring project was used by District scientists and management. Because the information from the existing monitoring plans, the interviews as well as the project data downloaded from DBHYDRO did not always agree, a process to reconcile individual working databases for each project was created. District staff approved the final project data sets prior to any statistical analyses being conducted.

Each of the above steps was found useful for reconciling the project data. Once all information was reconciled, the step that became critical for optimization was specifying, in detail, the uses of the monitoring data from each individual project. This step was essential to ensure the statistical approach to be used in optimizing was appropriate to the data use. The manner in which the monitoring data are, or will be used in the future drove the decision rules for the optimization. For all projects, the data uses were generally consistent and typically included:

1. Establishment of baseline for various water quality parameter concentrations;
2. Comparison of various water quality parameter concentrations to standards;
3. Estimation of the average change (primarily decreases) in various water quality parameter concentrations over a period of time;
4. Detection of a change in various water quality parameter concentrations as a result of an event or managed action;
5. Detection of the occurrence of high water quality parameter concentration events;

6. Detection of changes in water quality parameter gradients and estimation of the duration of those changes;
7. Estimation of the contribution a source from among a group of sources to a water quality parameter; and
8. Characterization of the relationship between various water quality parameter concentrations and land use/ land cover type.

The type of statistical metric that could be used to evaluate each of the data uses was identified. These metrics assumed that relatively simple statistical analyses could be applied and that performance criteria for each could be defined. However, as simple statistical procedures were conducted (i.e., time series) it became apparent that the data from the monitoring programs were much more complex (e.g., substantial autocorrelation and fixed seasonal effects) than could be handled by the simple statistical metrics planned, thus requiring a more complex statistical model. Given the time constraints of the effort, this realization that a complex statistical model would be necessary to conduct any optimization, raised the question as to whether all data uses for each individual project could be fully addressed. After further re-evaluation of the data uses for each project, it was determined the primary and most critical data use for each project was whether the current project monitoring design was sufficient to support trend evaluation and detection of changes in trends or movement towards a water quality standard. As a result, optimization of each of the projects focused primarily on these specific data uses. Where appropriate, the optimizations also focused on spatial evaluation. The remainder of this section discusses the general considerations for the optimization effort and the statistical procedures applied to optimize the programs for trend detection.

2.1 Project Considerations

During the optimization effort, issues common to several projects were identified. These include the use of autosamplers with grab sample backups, extensive parameter lists, use of *in-situ* data, and no clear data end user or uses of the data. Additionally, this effort was directed at concentration data only. This focus was problematic with many projects since loading appears to be a more critical issue than concentration. One other issue common among the projects was ability to monitor event-driven phenomena in the specific project area.

Autosamplers: Several of the projects undergoing optimization collect data via autosamplers and grab samples. Where a project uses autosamplers, many District staff wondered whether it would be necessary to continue the District protocol of collecting the grab sample as a backup for the autosampler in the case of autosampler failure. For these projects, Battelle included comparisons of autosampler versus grab sample data to assist the District in deciding whether grab samples could be eliminated or collected at a reduced frequency.

For the purposes of this effort, autosampler and grab sample data were compared using graphic representations and correlation/regression analyses. Consistent with several previous District comparisons, the analyses determined that these two sampling approaches generally give similar results. However, there are occasions when the autosamplers have unexplained low or high values. When examined concurrently with the autosampler data, grab sample data is often able to provide evidence as to whether the autosampler information should be accepted or rejected. Therefore, in cases where the optimization found that there is good correlation between autosampler and grab data, the grab sample may be eliminated, particularly if a mandate specifies sampling via autosampler only. However, the collection of grab samples as a backup must be carefully considered against project goals and the risk of data loss or uncertainty if no backup is available when uncertain or questionable results are encountered from autosampler information. Additionally, because the grab is collected when samples from the autosamplers are collected, the cost savings associated with eliminating the backup grab may be minimal (laboratory and data management costs only). Eliminating the backup grab samples may pose a considerable risk in cases where autosampler data is not available and the District must weigh the financial advantages to discontinuing backup grab samples with this risk. When faced with legal issues pertaining to water quality concentrations at specific locations, the costs saving associated with sacrificing a grab back up sample may not be a risk the District is willing to take.

Parameters: Another issue common to this optimization effort was that of the parameter list that was collected for any given project. For some projects, the list of parameters collected is extensive and, therefore whether all of these parameters are actually used by District scientists was evaluated. In some cases, parameters may have been requested by a scientist to address a certain issue, but when the issue was remedied, the parameter continued to be collected. During interviews with District staff, it was recommended that several parameters be dropped from all projects. These include chlorophyll b, chlorophyll c, carotenoids, and ammonia from autosamplers. Specific projects also recommended removal of several parameters that may fall into the category of those collected for a short period to address a specific issue but are not necessary any longer. For some projects, correlation analyses were conducted to determine if any of the parameters measured were highly related. Significant and strong correlations between parameters may allow for elimination of some parameters, however, District scientists will need to determine which parameters could be potentially eliminated based on their usefulness to District activities, and whether certain parameters are necessary for the interpretation of other parameter results. Additionally, the cost savings of eliminating some of these parameters may be minimal depending on laboratory methodologies (i.e., no additional costs to run various parameter species) as well as field sampling logistics (i.e., once you are at the location, collecting 10 ml extra water does not add additional costs).

In situ data: The use of *in situ* datasondes and the ability to attain meaningful results from these instruments was also an issue that District scientists suggested was problematic with several current monitoring projects. The current protocol for collecting *in situ* parameters such as dissolved oxygen (DO), specific conductivity (SCOND), pH, temperature and salinity during short (minutes) deployments does not provide the most scientifically robust nor desirable information to the end users. Rather, these parameters are more “usable” if the probes are deployed continuously for a longer period (e.g., up to 4 days) on a quarterly or seasonal basis. District staff scientists indicated that such deployments would enable measurement of diurnal cycles in the natural environment and provide the end users with more accurate characteristics of these parameters in the water bodies being monitored, particularly in areas experiencing low DO events or tidally driven salinity intrusions. This suggestion is considered appropriate and provides a scientifically based optimization that could result in a more efficient use of staff time for those calibrating and reviewing *in situ* data collections.

Concentration data: Many of the projects evaluated during this effort need to combine flow data (volume) with the water quality data (concentration or mass per unit volume) to estimate loading, export, or removal efficiencies to adequately address the data use. Ideally, optimization of projects concerned with these types of data uses would address the loading, not just the concentrations based data. As such several projects optimized during this effort identify the need to include flow and loading for effective final optimizations. It will also be essential that the District define and standardize methods for calculating loads. Additionally, the District must have access to accurate and detailed flow information to have confidence in the results.

Seasonality and autocorrelation: Due to the nature of the parameters being monitored (water quality parameters), much of the data are seasonal and highly autocorrelated (e.g., data from one time period not independent of the preceding measurement). Water quality data from South Florida generally falls into two seasons, wet and dry, and review of time series plots depict regular patterns in the data that correspond to the wet and dry seasons. Moreover, water quality data often exhibited high degrees of correlation from one sampling period to the next. These two issues, and the fact that data that are not normal distributed, required more sophisticated analysis techniques including the consideration of non-parametric statistical approaches.

Identifying data uses: The key for this optimization effort was gaining a clear understanding of who was using the data and information generated from an individual monitoring project and how was that information being used. For most of the projects, the data end users could be identified. How the data were used to answer scientific questions, inform management and operations etc. could was also characterized. For project RAIN, which has been ongoing for 25+ years, data end users and data uses could not be identified. This specific project has been brought before the Technical Oversight Committee (TOC) in the past where it was recommended it be dropped. However, the project is still being conducted and multiple contractors collect a suite of parameters from several stations. Additionally, District scientists suggested that the quality of this information is suspect due to problems with environmental contamination of the samples. No statistical analyses were conducted for this project because specific uses for the information could not be identified. It is recommended that this project be eliminated from the current District monitoring efforts.

Event-driven phenomena: One theme that was present in many of the projects was the ability to address event-driven phenomena. The best current technology to achieve the collection of event driven data is through flow proportional automatic sampling. There are several constraints associated with this type of sampling, including costs for instrumentation and the holding times for parameters other than nutrients. Clear definitions of what constitutes an event were not apparent during the optimization, thus this effort did not focus on the event-driven sampling design considerations. One way to get a better handle on this problem is for the District to develop an event driven operational plan that would ensure data appropriate for event monitoring is obtained. Such a plan should include a precise statement of the problem and question(s) being addressed, an evaluation of alternative sample collection techniques, and sampling frequency necessary to address the issue.

2.2 Statistical Methods

As indicated previously, initial considerations of relatively straight forward statistical approaches to optimization could not be applied due to the complexity of the data which included a large degree of seasonality, high autocorrelation, and non-normal data distribution. Any type of parametric procedure requires the data to be normally distributed and no mathematical transformations could be applied to render these data more normal or address the issues of seasonality and autocorrelation.

Because the District often employs non-parametric statistical approaches to evaluate data, non-parametric approaches were used to the extent possible in all statistical analyses conducted for the individual projects. To address the principal data end use of detecting trends and changes from those trends, non-parametric analyses combined with the use of statistical modeling were used to develop the procedure for trend detection for each of the projects. This procedure evaluated the current monitoring design and allowed evaluation of series of alternative sampling frequency designs. The procedure was developed in SAS and relies on methods from the Seasonal Kendall Tau trend analysis developed by Reckhow *et al.,* 1993. The SAS code and step by step instructions for this procedure was supplied to the District under separate cover (Rust 2005) as a tool that District scientists may use for future optimizations or when determining new monitoring programs. The individual statistical tests used in the procedures are described briefly below. Details of the procedure are provided in Section 6.

In addition, time series plots and box plots were used to initially evaluate the data from any given project. Several other statistical procedures were also used in the various optimizations to examine the data on a temporal basis, on a spatial basis, and for the parameter. Brief descriptions of these analyses are summarized below.

Trend power analysis
Monte Carlo simulations were performed using the Seasonal Kendall Tau Test for Trend which estimates the power to detect a trend for a given water quality parameter. The Seasonal Kendall Tau trend analysis procedure (Reckhow et al. 1993) was used to simulate time series data set (individual sites or groups of sites/stations) to obtain a point estimate of the slope vs. time for the log-transformed water quality parameter values for the current monitoring design and under alternative sampling strategies. The tests are run under a power of 0.80 ($\beta = 0.2$) and $p = 0.05$. A 20% change in slope of any given parameter over a five year time period was used as a target change for detection. Key outputs are the annual percent change (APC) that the monitoring scenario is able to detect.

Nonparametric Sign Test
The sign test (Zar 1984) was used in a simulation experiment to assess the power to detect changes in the distribution for a particular parameter of interest from a target value or, in the absence of a target value, a baseline condition (i.e. the long term median). This simulation trial was established to assess the minimum difference in median value that can be detected under the current sampling scheme and identify the sample size requirements necessary to detect a change to either a target value or a change of 20% in the long term median. The 20% criterion was established *a priori* as a change in median signifying a significant shift in the parameter distribution. A Monte Carlo approach was used to create the dataset for simulation. The Sign Test Monte Carlo simulations did not account for serial auto-correlation which can be present in monitoring data. The presence of significant auto correlation, if not accounted for, can yield unrealistically optimistic assessments of the sample size necessary to detect changes. However, from a regulatory perspective, auto-correlation is usually not considered when assessing whether or not a water body is meeting or exceeding a given water quality target (e.g., Impaired Waters Rule F.A.C. 62-303.320).

To determine the effect size (the magnitude of difference that could be detected) the distribution was shifted by multiplying each value by a constant (e.g. median value*0.20+median value). Experimental trials were created by re-sampling from the shifted distribution and altering the number of observations for each experiment to assess a range of sample sizes corresponding to potential sampling schemes for the project. For each experiment in the simulation, 400 replicate trails were performed in which an experimental sample of data was selected, the difference between each value and the long term median was established and a sign (+ or -) was assigned to each record to indicate if the difference was positive of negative. The proportion of positive signs was then assessed to quantify whether the proportion of positive signs was significantly different from 0.50 (the expectation under the null hypothesis). Results for each trial were tallied and if 95% of the results were statistically significant using an alpha level of 0.05, the difference was deemed a detectable difference.

Spearman's Rank Correlation
Spearman's rank correlation analysis is a non-parametric version of the Pearson Product Moment correlation analysis. This approach was used to evaluate the correlations between sampling stations (for specific parameters) as well as to evaluate correlations between all parameters sampled for a given project.

Wilkoxan Rank Sum Test
The Wilcoxan rank sum test was used in several instances to examine similarities and degree of covariance between specific stations.

Principal Component Analysis
PCA analysis (Hatcher and Stepanski 1994) is a data reduction technique used to reduce a large number of variables which may be highly correlated into PCA axes which have three important features: Each resulting PCA axis is uncorrelated with the others; it orders the PCA axis so that those accounting for the

largest variation come first, and it eliminates components which contribute little to the overall variation in the data. In this study PCA was used to identify stations that were functionally similar with respect to their variation over time for the parameters of interest. The PCA analysis was used to identify stations that may be providing redundant information to assessing variation of a particular parameter of interest within the sampling program.

Optimization approach summary
In summary, several types of optimizations were conducted across the projects and whether or not they were performed and the tests used to perform them vary by project. The types of optimizations can be characterized as data use, parameters, station redundancies and spatial grouping, temporal trends, and special considerations. Table 2 summarizes the types of optimizations conducted for each of the 18 projects optimized as primary and secondary optimizations. The type of optimization conducted is listed. For these projects, only a portion of the data uses identified in the approved project summary could be addressed by this optimization.

Table 2. Summary of Optimization Types Conducted versus by Project. NA = not applicable; N not performed.

Project	Temporal	Spatial	Parameter	Other
BISC	Trend Power	Cluster	Tech review	Interlaboratory
BRM	Trend Power Sign Test	N	N	
CAMB	Trend Power	N	N	Autosamplers
CCWQ	Sign Test	PCA Spearman's	N	
CESWQ	Trend Power Sign test	Spearman's	N	
CR	Trend Power Sign test	Spearman's	N	
IRL	Trend Power Sign test	N	N	
KREA	Trend Power	NA	N	
LKR	Trend Power	NA	N	
OLIT	Trend Power	Spearman's	Spearman's Tech review	
RAIN	NA	NA	NA	
SE	Trend Power Sign test	Spearman's	Tech review	
SEMI	Trend Power	Spearman's	Tech review	Autosampler
ST1W	N	F test	Tech review	
TCNS	Trend Power	NA	N	
WQM	Trend Power Sign test	Spearman's Wilcoxon	N	Autosampler
V	Trend Power	NA	N	
Y	Trend Power	Spearman's	Spearman's Tech review	

3. SUMMARY OF CURRENT MONITORING PLANS

This section briefly considers the characteristics of the project optimized. The District funds and conducts the sampling efforts for all but one of the projects described.

BISC: The District instituted the Biscayne Bay (BISC) Monitoring program in 1978 and updates to the program occurred in 1995. Sampling occurs in all of Biscayne Bay from the Broward County line to U.S. Highway 1 at Key Largo, The tributaries/canals leading into the Bay are also sampled as part of this monitoring program. All sampling stations are either Type II or Type III mandated. Two agencies currently conduct the sampling of Biscayne Bay for the District, Miami Dade Department of Environmental Resource Management (DERM) and Florida International University (FIU) under funding from the state of Florida and District, respectively. DERM samples 113 stations (41 in the Bay and 72 in canals) for 22 parameters. Sampling frequency is monthly or bimonthly for nutrients and quarterly for metals. Thirty-five locations are sampled by FIU on a monthly basis for 16 parameters, several of which are the same as those sampled by DERM. The BISC data are used for a number of different purposes including evaluating spatial and temporal trends, identifying hotspots for select parameters, development of stormwater improvement programs, development of non-degradation criteria, and development of freshwater response relationships. In the future, data from BISC will be necessary for evaluating the effectiveness of specific CERP projects, evaluating minimum flow criteria and long-term monitoring for RECOVER.

A key question for using data from multiple agency programs is data comparability. To evaluate this for BISC, a comparison of the data generated by the two programs was completed. The evaluation consisted of determining stations that were closely located geographically, then comparing a set of parameters using box plots, parameter by station plots, and time series plots to determine which parameters were comparable or not.

BRM: The Brighton Reservation Monitoring Program (BRM) was initiated in 2002 to address concerns of spiked phosphorus concentrations observed on the Reservation that did not appear to be related to internal practices. Four stations that are Type II mandated are sampled for Project BRM; however, data from six, Type I mandated stations from Project X must be evaluated with the BRM stations to address the phosphorus concentration issues. The four BRM stations are sampled weekly for nitrogen (TKN and NOX) and phosphorus only. These data are used in the Lake Okeechobee Watershed Assessment report and are included in reports to the Seminole Tribe.

CAMB: The Conservation Area Inflows and Outflows monitoring program was initiated in 1977. CAMB was created to comply with water quality monitoring requirements of the Everglades National Park Memorandum of Agreement between the National Park Service, the District and the Corps. The program is essentially a selection of stations that are being monitored to respond to various mandates. They have been grouped into this program based on their locations throughout the water conservation areas. Thirty-seven stations are sampled for this project. Several stations are sampled under multiple mandates and may, therefore, be both a Type I and Type II mandated station. Twenty-five parameters are measured for this project, but they vary by station. Frequency also varies by station. Parameters and the frequency with which those parameters must be monitored for Type I stations are specified in the mandates. The data from CAMB are used to determine the effectiveness of basin management plans to reduce nutrient loading to the water conservation areas and to establish nutrient budgets for the water conservation areas. Additionally, this monitoring is used to quantify the effects of inflows on the ecology of marshes.

CCWQ: The Collier County Water Quality Monitoring program (CCWQ) was instituted in 2000 to support the District's commitment to provide data to better address water quality issues in Southwest Florida. Forty-eight stations are currently sampled under CCWQ. Subsets of the stations are Type I mandated under the Prairie Canal permit and Corkscrew Swamp permit with DEP. The stations located in Picayune Strand, a CERP Acceler8 project, are Type II and all others are Type III. Twenty-five parameters are sampled for CCWQ on a monthly basis and fourteen parameters are sampled quarterly. The data from CCWQ are used to evaluate baseline conditions and look at trends in water quality parameters in the Big Cypress basin watershed and adjacent coastal waters of Collier County. Additionally, these data will be necessary for the Southwest FL Feasibility Study, the District's Water Supply Plan for the Reservations and various CERP projects.

CESWQ: The Caloosahatchee Estuary Water Quality Monitoring program (CESWQ) originally began in 1998 but was re-designed in 2002. CESWQ consists of regular monthly sampling as well as event-driven monitoring. Monthly sampling for 21 parameters is conducted at 4 fixed stations and 5 randomly-selected stations to better understand trends in these parameters in the Caloosahatchee and receiving estuaries. The event-based sampling effort is conducted to help quantify the effects of freshwater releases from Lake Okeechobee to the Caloosahatchee River Estuary. Event-based sampling is conducted at 11 stations for six parameters. Nitrogen and phosphorus are only measured at one station during the event-based efforts. Data from CESWQ are critical to many District reports and models, as well as District operations. These data will be needed for the C-43 basin CERP Acceler8 project as well as the RECOVER Monitoring and Assessment Plan.

CR: The Caloosahatchee River Monitoring Program (CR) was initiated in 1979 to implement long-term monitoring to better evaluate both short-term and long-term trends in several water quality parameters in the Caloosahatchee River. Four stations are sampled for this program and all are Type II mandated. Twenty-two parameters are measured at each station on a bi-monthly basis. Data from CR are critical to many District reports and models, as well as District operations. These data will be needed for the C-43 basin CERP Acceler8 project as well as the RECOVER Monitoring and Assessment Plan.

IRL: The Indian River Lagoon Monitoring Program was started in 1988 and has recently undergone a re-design to make the program as efficient as possible. Water quality monitoring in the lagoon is necessary to document short- and long-term trends in several water quality parameters. Additionally, monitoring of the IRL is necessary to evaluate the link between water quality and seagrass health. Twenty-one stations are sampled seven times per year for twenty parameters. All sampling locations are in association with seagrass beds and are Type II mandated. The data from this program are used in numerous District reports and modeling activities. District operations also use the data to evaluate the lagoons response to releases from Lake Okeechobee. Data from the IRL project will be critical to the North Palm Beach County CERP projects and the Indian River Lagoon South CERP project. Many of the monitoring stations from the IRL project will also be included in the RECOVER Monitoring and Assessment Plan.

KREA: The Kissimmee River Eutrophication and Abatement Project (KREA) was initiated in 1986 to provide baseline and assessment data for watershed restoration and enhancement projects in the Kissimmee River basin. Sampling occurs along many of the tributaries that drain dairy and agricultural areas. Twenty-three sampling stations, which are Type II mandated, are currently sampled for KREA. Ten additional stations have been sampled for this project in the past; however, they are now incorporated under the Lake Okeechobee Watershed Assessment (LOWA) program. Sampling is conducted bi-weekly for 11 water quality parameters at 13 stations and conducted monthly at 10 stations for 20 parameters. Ions are collected quarterly at 10 stations. KREA data are used in many District reports and Lake Okeechobee watershed modeling activities. These data will also be necessary for the CERP watershed critical project, the Taylor Creek Nubbin Slough Stormwater Treatment Area (STA).

LKR: The Lower Kissimmee River Monitoring program (LKR) was initiated in 1987 to assess tributary, basin loading and concentration inputs of phosphorus to the Kissimmee River and Lake Okeechobee, as well as to evaluate temporal trends in phosphorus in these areas. Sampling is conducted at seven stations. Three stations are Type I mandated (they are considered Type I for Project X) and the remaining stations are Type II mandated. Phosphorus is the only parameter measured and it is collected weekly via autosampler. The data from Project LKR are used in several District reports and Lake Okeechobee watershed modeling activities. Several stations sampled under LKR may also be used for long-term monitoring for the RECOVER Monitoring and Assessment Plan.

OLIT: The Lake Okeechobee Littoral Zone Monitoring program (OLIT) was initiated in 1996 to look specifically at the littoral zone (marsh area) of the lake disconnected from the lake proper. Twelve sampling stations (all Type II mandated) representing three geographic areas within the lake's western littoral zone are sampled monthly for 23 parameters. The data are used to identify short- and long-term trends in various water quality parameters and to determine the effectiveness of basin management practices in reducing nutrient loads to the lake. Data from OLIT are used in a number of District reports and models as well as operations. OLIT data are used with data from Y to compare the littoral and limnetic zones of the lake and monitor algal bloom conditions. In the future, data from OLIT will likely be used to monitor impacts from CERP activities. Additionally, the OLIT stations may be monitored for the RECOVER Monitoring and Assessment Plan.

RAIN: The Rain and Atmospheric Deposition program (RAIN) started in 1974 to evaluate nutrient concentrations in wet atmospheric deposition and determine the resulting nutrient loads to the south Florida ecosystem from this medium. Currently, five stations are monitored weekly for fifteen parameters. The quality of the data is suspect and collectors have been modified with hoped of improving data quality. The District has been unable to determine who uses this information and whether it is used in any District reporting requirements.

SE: The St. Lucie Estuary monitoring program was initiated in 1989 to evaluate both short- and long-term trends in water quality parameters in the estuary. The data from this program are also used to determine the effects of freshwater releases upon seagrasses, oyster beds and macroinvertebrates that inhabit the system. Thirteen stations, all Type II mandated, are sampled monthly for twenty parameters. Data from the SE program are critical to a number of District operations and reports. Like the IRL program, these data are used to evaluate the impacts of freshwater releases from Lake Okeechobee on the estuary. Along with data from the IRL, data from SE will also be critical to the North Palm Beach County CERP projects and the Indian River Lagoon South CERP project. Many of the monitoring stations from the SE project will also be included in the RECOVER Monitoring and Assessment Plan.

SEMI: The Seminole Reservation Water Quality Monitoring program (SEMI) was initiated in 1996 to satisfy the requirements of the agreement between the SFWMD and the Seminole Tribe of FL. The goal of the project is to determine the quality of water in terms of phosphorus that flows into the Big Cypress Indian Reservation. The stations selected for monitoring are Type II mandated and were stipulated in the agreement. However, six of the eight stations are also considered Type I mandate under the EAA Rule or Settlement agreement. Total phosphorus is measured weekly at all stations. The data for SEMI are used in many District reports and may be necessary for CERP, particularly the L28 levee system projects.

ST1W: Stormwater Treatment Area 1 West (ST1W) was constructed in two phases, with the initial phase of construction (Cells 1-4) completed in 1994. This initial phase, often referred to as the Everglades Nutrient Removal Project, was constructed to begin to evaluate the effectiveness of phosphorus removal within a large constructed wetland in South FL. ST1W was subsequently expanded in 2000 with the completion of Cell 5. The primary purpose of the ST1W monitoring program is to respond to the Everglades Forever Act mandate which requires the annual reporting of phosphorus into and out of each cell within ST1W. The main focus of the monitoring is to determine the long-term phosphorus removal performance of each cell, both independently as well as within a flow path. Additionally, data from this program are used to calibrate a dynamic operational model and provide direction to further optimize the STA to reach a phosphorus criterion of 10 ug-P/l. The data are also used for management and operational decisions and the impacts these may have on STA performance.

Project STIW has undergone several internal evaluations and changes in response to permit discussions and has had an evolving monitoring design. Due to the changing nature of the monitoring design, questions regarding whether the spatial adequacy of water quality stations used to sample outflow locations to ensure the estimated phosphorous export (outflow) are accurately described remain. Statistical evaluation of the available data from two levee sites found that the outflow from a cell can be variable and is not always the same across the cell boundary. To ensure the uncertainties are better understood, a short-term autosampler study (6 months to 1 year) that samples cell boundaries at a more highly resolved spatial scale and over a range of outflow conditions is recommended.

TCNS: The Taylor Creek Nubbin Slough Monitoring program (TCNS) was initiated in 1979 to generate baseline and assessment data for watershed restoration and enhancement projects. The TCNS project lies within an area characterized by beef and intensive dairy cattle operations. Best Management Practices have been implemented in this watershed for the Works of the District Program as well as the Dairy Rule and Rural Clean Waters Program. The data from TCNS are used to evaluate the efficacy of BMPs for reducing phosphorus in surface water discharges from diaries, evaluating phosphorus contributions from each tributary, estimating phosphorus loads leaving the basins and identifying high episodic events and locating source areas. Sampling for TCNS is conducted biweekly at 11 stations for 11 parameters and monthly at 3 stations for 6 parameters. All stations are Type II mandate. Ten additional stations that have been sampled as part of TCNS in the past are currently sampled under LOWA. The data from KREA are used in several District reports and Lake Okeechobee watershed modeling activities.

V: The Kissimmee River Structures Monitoring program (V) began in 1973 to assess tributary, basin loading and concentration inputs to the Kissimmee River and Lake Okeechobee, as well as identify trends in various water quality parameters over time. Five stations are sampled for project V and these correspond to five of the seven locations sampled under LKR. However, unlike LKR which uses autosamplers at these locations, sampling for V consists of grabs. Four of the five stations sampled under V are Type II mandated, whereas one station sampled under V is Type I mandated under Project X. The five stations are sampled biweekly for 23 parameters and quarterly for three parameters. Like the LKR data, V data are used in several District reports and Lake Okeechobee watershed modeling activities. Several stations sampled under V may also be used for long-term monitoring for the RECOVER Monitoring and Assessment Plan.

WQM: The Lake Worth and West Palm Beach Water Quality Monitoring Network program (WQM) was initiated in 2002. This program serves as one of several projects to implement a comprehensive research and monitoring program called for by the Lake Okeechobee Technical Advisory Committee. The monitoring stations for this program were established to identify seasonal and discharge related water quality trends and determine loadings to the Indian River Lagoon, St. Lucie Estuary, Loxahatchee River and Lake Worth Lagoon. Eleven stations, all Type II mandated, spanning from St. Lucie County, Martin County and Palm Beach County are sampled monthly or quarterly for 22 – 26 water quality parameters depending on location. Stations within St. Lucie and Martin County also are fitted with autosamplers for weekly measurements of nitrogen (NOX and TKN) and phosphorus. Data from WQM are critical to a number of District reports, models and operations. District operations uses data from WQM, along with that from SE and IRL, to evaluate the impact of releases on freshwater on the estuarine systems. Data from WQM, along with SE and IRL, will also be critical to the North Palm Beach County CERP projects and the Indian River Lagoon South CERP project. Many of the monitoring stations from WQM will also be included in the RECOVER Monitoring and Assessment Plan.

Y: The Lake Okeechobee In-lake North and In-lake South Monitoring programs (YNRG/YSRG or Y) measure water quality parameters in the limnetic area of Lake Okeechobee as opposed to the littoral zone that is monitored under OLIT. This program began in 1972 with the principal goal of establishing baseline water quality parameter concentrations and determining spatial and temporal trends within the lake. Today, the data from Y are used for numerous purposes including assessing the impacts of District operations, changes in water quality due to basin management strategies, verifying water quality models, evaluating the differences between the limnetic and littoral zones, monitoring potential for algal blooms and establishing nutrient budgets for the lake. Twenty-seven stations, all of which are Type II mandated are sampled on a monthly basis for 24 parameters. Four additional parameters are sampled from these stations on a quarterly basis. In the future, data from Y, along with OLIT, will likely be used to monitor impacts from CERP activities. Additionally, the Y stations may be monitored for the RECOVER Monitoring and Assessment Plan.

4. RECOMMENDATIONS FOR MONITORING PLAN OPTIMIZATIONS

This section of the report addresses four areas: 1) the primary optimization recommendations for each individual project; 2) a compilation of overarching observations and recommendations derived from the optimization effort, 3) a summary of the benefits that may be gained from implementation of the optimization recommendations, and 4) partnering opportunities.

4.1. Project-specific Optimization Recommendations

The primary recommendations for optimization of individual projects are summarized in Table 3. The table addresses recommendations regarding changes or modifications to the stations sampled, sample collection frequency, and parameters measured. Detailed discussion of the results of each optimization and corresponding recommendations can be found in the individual Project Summary Updates are included in Attachment 2.

The period of record for BRM, CCWQ, CESWQ, and IRL was generally too short to conduct adequate statistical optimizations. Several additional years of information would enhance the optimization effort. Potential revisions to the number of sites/stations sampled were identified for eight projects (BISC, CAMB, CCWQ, CESWQ, CR, OLIT, WQM, and Y), although the number of potentially redundant stations was limited to one or two in most of the projects. The open water-bodies such as Lake Okeechobee, Biscayne Bay, the Indian River Lagoon, Caloosahatchee Estuary, and St. Lucie Estuary had greater potential for modifying the monitoring program either by removing stations from the sampling plan or by modifying the fixed station sampling strategy to a stratified random approach. A possibility identified for reducing the sampling effort associated with projects that focus on source identification and nutrient loading control (e.g., CR, KREA, LKR, TCNS, and V) was to make use of a regional approach which would require spatially averaging data rather than using individual station data to make judgments regarding trends. Implementation of this strategy would require the District to reevaluate its site by site monitoring strategy and potentially require revision of the mandate requirements that direct the monitoring. Pursuing this optimization strategy would also require a clear definition of the data aggregation, the endpoint and decision metrics. The power of data aggregations to address mandate requirements was not pursued under this optimization as it requires policy discussions within the District. The only project where additional stations were recommended to improve the ability to detect trends was salinity in BISC.

The optimization of sampling frequency was problematic within each project given the high temporal variability and fixed seasonal effects resident within most projects. There was also substantial autocorrelation within the time series data for most parameters within a project which limits the ability to obtain truly independent samples. Temporal optimizations depended on the parameter evaluated and the magnitude of the annual percent change in the trend that is desirable. The target 20% annual change over a five year period could not be met in many cases, was exceeded greatly in some cases, and met under the current or more frequent sampling scenarios in other cases. However, most project monitoring designs appeared to be reasonably optimal if an annual percent change of 10 to 30 percent over a five year period is acceptable. Projects that optimization results suggested could be less frequently sampled include BISC (for laboratory based parameters such as TPO4 and TOTN), CAMB (selected sites, sampling methods, and parameters), LKR (selected stations), SE (some stations), and SEMI (grab samples and autosamplers at some stations). Increased sample collection frequency was identified as a way to improve the annual percent change detected in BISC (salinity only), turbidity in CAMB at one site, IRL for parameters other than TPO4 and PAR, SE for some stations and parameters, and TPO4 in TCNS at some stations. The frequency optimization for other projects and parameters depends greatly on the annual percent change that must be detected as well as the time period required to detect a trend. These considerations require additional considerations within each project and the District in general.

In spite of this current extensive review and evaluation there remain several questions and additional analysis to fully optimize each of the projects. These include defining, in detail, the acceptable amount of change in trend that is detectable at a specified statistical power. This is necessary in part due to the uncertainty in the endpoints (amount of change expected /required) the project's targets and also the metrics used to evaluate the data (e.g., individual sites versus data aggregation; annual versus seasonal versus monthly versus daily compliance, etc).

Table 3. Summary of Project Optimization Recommendations.

Project	Spatial	Temporal	Parameters
BISC	Drop or relocate one of the stations in each of the fifteen contiguous station pairs that provide redundant physical (salinity, temperature) information. Consider having only one program monitor co-located stations presently sampled by the two programs and shared data (assuming data comparability requirements are met). Consider redistribution of redundant and co-located stations to fill spatial gaps in the along shore region in central Biscayne Bay to provide better ability to detect long-term change across the hydrographic gradients within this region of the Bay. Consider relocating offshore DERM stations BB35, BB37, BB38, and BB44 closer to shore to capture more of the near shore variability associated with C100 canal; use FIU data to address offshore DERM monitoring requirements. Consider redeployment of non-MAP critical DERM of stations in Northern Biscayne Bay that provide redundant data to other near coastal areas or drop these from the program. Evaluate whether stations located in	Consider less frequent sampling for all parameters except salinity. Statistical results suggest the present monitoring programs are not robust enough to detect annual percent changes in salinity to within 100 to 400 percent. Detection of smaller changes will require higher frequency and likely more spatial stations. Depending on the desired annual detectable change, the other parameters could be sampled less frequently than at present, especially TOTN and TPO4. Since these are laboratory based measures and other parameters such as chlorophyll and dissolved oxygen can be obtained *in situ*, downward revision of the frequency these laboratory based parameters should be considered.	Consider modifying or dropping TCMF and FCMF from offshore water monitoring except near shellfish resource areas. Monitoring for these parameters should be linked to climatologically or other events (water releases to determine if an event response monitoring program is more appropriate for these in offshore waters rather than a routine fixed frequency program). Consider dropping measurement color in offshore ocean waters unless associated with seagrass recovery or areas. Consider dropping metals measurements in offshore waters unless there are clear transport and exposure issues that bring contaminated water offshore and over critical resource areas. An event based monitoring program may be a more appropriate approach to metals monitoring. The MAP (RECOVER 2004) suggests adding silicate, TKN, reactive PO4 (*aka* OPO4), more metals and chlorophyll sampling to the Biscayne Bay monitoring program and at higher frequency. The statistical evaluation does not support addition of more total nitrogen sampling either as TKN or TOTN. If any total nitrogen measurements are to be added to the DERM program TOTN is more appropriate to enable the two programs to produce more comparable total nitrogen data. The measurement of photosynthetically active

Project	Spatial	Temporal	Parameters
	Southern Biscayne Bay that are routinely monitored under the Florida Bay Monitoring Program (FLAB) should be included in BISC. Consider relocation of the FIU stations BISC129 to BISC134 from northern Biscayne Bay to the along shore regions of the west central Bay or in southern Biscayne Bay in Card or Barnes Sound.		radiation (PARK [197]) of PARK should be restricted to those areas most likely to have concerns regarding reestablishment or maintenance of seagrass. Retention of parameters that support water quality modeling and understanding ecological processes that may change as water flow to the coast is modified under CERP is important and the mark of quality monitoring programs that strive to understand the whys and wherefores of observed changes.
BRM	Period of record is too short to adequately optimize the project. Continue monitoring at current spatial design to build the baseline for later optimization.	Period of record is too short to adequately optimize monitoring. Continue monitoring at current frequency to build the baseline for later optimization.	No changes recommended.
CAMB	Stations S6 and S10D may be redundant for several parameters. Consider dropping one of these stations. Consider using load data to optimize the project by defining acceptable uncertainty and annual percent change necessary to meet mandate requirements.	**TKN**: <u>Grab</u> samples at sites L3BRS, S140, S190, S5A, and USSO could be reduced to monthly. <u>Auto</u> sampling at sites S6, USSO could be reduced to monthly and at S5A reduce to bi-weekly. **TPO4**: Grab samples at site L3BRS could be increased to weekly. Those from sites S140, S5A, USSO decreased to monthly and Site S6 decrease to bi-weekly. **Auto** samples from sites S140, S190, and USSO could be decreased to monthly. **NOX**: Retain <u>Grab</u> sample frequency at site S6 at current weekly rate. Consider reducing <u>auto</u> samples from site USSO to bi-weekly.	Consider eliminating either the auto or grab sample for parameters measured by both auto and grab sampling methods.

Project	Spatial	Temporal	Parameters
		Turbidity: Consider increasing sampling at L3BRS to weekly. Retain frequency at sites S140, S5A, S6, USSO and S190. **CL**: Consider reducing sampling frequency at sites L3BRS, S190, S5A, and USSO to monthly.	
CCWQ	There are two areas (western Cocohatchee Canal [5 sites] and confluence of the Tamiami Canal and Henderson Creek Canal, [2 sites]) where Type I and III stations are in relatively close proximity and not synoptically sampled. Potentially redundant station pairs: 1) Type III Sites BC20 and BC21; 2) Type III Sites BC9 and BC10 appear to statistically redundant although they sample different sources; 3) Type III sites ECOCORIV, COCAT41, and Type I Site BC15	Period of record is too short to adequately optimize monitoring. Continue monitoring at current frequency to build the baseline for later optimization.	Specific goals/targets and uses for data must be identified to determine optimal parameters to sample.
CESWQ	Consider changing all stations except station CES01 to a stratified random sampling design. Retain CES01 as a fixed location station Consider using load data to optimize the project by defining acceptable uncertainty and annual percent change necessary to meet mandate requirements.	Basin alternations are not complete and period of record is too short to optimize monitoring. Continue monitoring at current frequency to build the baseline for later optimization.	No changes recommended.
CR	Consider concentrating monitoring	Sampling frequency will depend on	Discontinue TSS at all stations except S78 and

Project	Spatial	Temporal	Parameters
	stations in the western portion of the Basin at Stations S78 and S79. Consider using load data to optimize the project and defining acceptable uncertainty and annual percent change necessary to meet mandate requirement.	needs for nutrient loading calculation. Determine acceptable load calculation uncertainty.	S79. Consider adding NOX.
IRL	Evaluate the water quality criterion and endpoints established for each parameter of interest for appropriateness and effect on the optimal number of sample locations and frequency. Reevaluate the manner in which the mandate metrics are applied. Considering a meta-analysis or station aggregation approach to assess basin-wide trends in addition to using the median condition from the thirteen stations.	Project sampling is at the minimum frequency necessary to detect trends in water quality at most stations. Sampling frequencies appear adequate to detect trends for TPO4 and PAR. Higher frequency sampling is required to detect trends in other parameters. A longer time series is required to complete the temporal optimization.	Discontinue Chlorophyll b, c, and carotenoids measurements.
KREA	Consider using alternative numbers of sampling stations to detect trends and changes from those trends and running additional power analyses. Consider using load data to optimize the project and defining acceptable uncertainty and annual percent change necessary to meet mandate requirements.	Increasing or decreasing sampling frequency does not greatly improve the ability to reach the target APC. Investigating methods to aggregate data and methods (likely use of load, not concentration data) to analyze parameters concentration data to better explain the systematic variations over time.	No changes recommended.
LKR	Consider using alternative numbers of sampling stations to detect trends and changes from those trends and running additional power analyses.	Consider reducing sampling frequency at S65, S65A, S65C, and S65E to 24 samples per year (twice monthly).	No changes recommended.

Project	Spatial	Temporal	Parameters
	Consider using load data to optimize the project by defining acceptable uncertainty and annual percent change necessary to meet mandate requirements.	No changes are recommended for S154. Investigate more sophisticated methods (likely use of load, not concentration data) for analyzing TPO4 at this location to better explain systematic variations over time.	
OLIT	Preliminary spatial analyses suggest current numbers of stations in the North Littoral Zone and Fisheating Bay should be maintained. Several South littoral zone stations appear to be redundant for some, but not all, parameters. Additional spatial analysis such as Principal Component Analysis (PCA) coupled with additional power analyses is recommended to evaluate the ability to detect trends and changes from those trends using alternative numbers of sampling stations.	Increasing sampling frequency does not greatly improve the ability to meet a 20% APC target. Maintain current frequency if target APC is sufficient for mandate needs. Investigate additional data aggregation and statistical methods to evaluate trends and frequencies.	Chlorophyll b and c, carotenoids, color, magnesium, sodium, and VSS do not appear to be used. These measurements should be discontinued.
RAIN	Drop this project from the District's monitoring plan as no data users could be identified.		
SE	Consider changing reporting requirements to stratum rather by station to reduce sampling effort. Consider a stratified random sampling plan if a stratum approach is used. Consider using load data to optimize the project by defining acceptable uncertainty and annual percent change necessary to meet mandate	Reduce sampling frequency at some fixed stations. Water quality trend detection may require weekly sampling under the present design. A strata approach may allow lower sampling frequency due to high individual station correlations.	Discontinue Chlorophyll b, c, and carotenoids

Project	Spatial	Temporal	Parameters
	requirements.		
SEMI	Maintain current spatial design for monitoring. Consider using load data to optimize the project by defining acceptable uncertainty and annual percent change necessary to meet mandate requirements	Consider reducing grab sampling frequency at sites G357, G404, G409 and WWeir from weekly to bi-weekly. Consider reducing autosampler sampling frequency from weekly to bi-weekly at G357 and USSO.	No changes recommended
ST1W	Study is continually optimized by the District through data evaluations and program decisions. A special study to determine optimal sampling sites along levees is needed to address spatial variability in water leaving Cell 5b Additional optimization efforts must use load data to be effective.	Temporal optimization must include load data. Optimal temporal sampling requires paired input and output sampling over time to account for lag time in input output load calculations.	No changes recommended.
TCNS	Consider using alternative numbers of sampling stations to detect trends and changes from those trends and running additional power analyses Consider using load data to optimize the project by defining acceptable uncertainty and annual percent change necessary to meet mandate requirements.	Consider increasing sampling frequency for TPO4 to weekly at stations 201, 204, 212, 233 and 249. Investigate more sophisticated methods (likely use of load, not concentration data) for analyzing TPO4 at other locations to better explain systematic variations over time.	No changes recommended.
WQM	Consider eliminating either Station C18S92 or C18S46.	Frequency sufficient for the most part to detect 25% annual changes. Consider if this level of change detection is sufficient for the project.	Use of autosamplers is preferred sampling technique if loading estimates are the primary data use.

Project	Spatial	Temporal	Parameters
V	Consider using alternative numbers of sampling stations to detect trends and changes from those trends and running additional power analyses. Consider using load data to optimize the project by defining acceptable uncertainty and annual percent change necessary to meet mandate requirements.	Maintain current sampling frequency at all stations until more sophisticated methods (e.g., use of load, not concentration data) for analyzing concentration data to better explain systematic variations over time.	No change, but the importance of detecting detect trends in parameters other than TPO4 should be evaluated.
Y	Preliminary spatial analyses suggest that stations within the specific regions of the lake may be redundant for some parameters. Additional spatial analysis such as Principal Component Analysis (PCA) coupled with additional power analyses is recommended to evaluate the ability to detect trends and changes from those trends using alternative numbers of sampling stations.	Maintain current frequency However, current sampling does reach the 20% APC target for some parameters and stations. When the 20% APC target is not attained, increasing the sampling frequency does not always improve the ability to meet this target. Investigate additional methods to evaluate concentration data.	Chlorophyll b and c, carotenoids, color, magnesium, sodium, and VSS do not appear to be used These measurements should be discontinued.

4.2 Overarching Optimization Observations and Recommendations

The individual optimization findings and recommendations coalesced around several general themes relevant to the District's water quality optimization initiative. These include 1) expanding the optimization for several projects from concentrations based data to loading which incorporates flow data (i.e., optimize on loading rather than concentration), 2) more precisely defining the level of change and period over which the change should be detected so as to improve further temporal optimizations (the fact that the five year period used for the trend evolution is too short to detect changes against high variability, fixed seasonal effects, and autocorrelation within the data, 3) clarifying data uses and data use matrices better (improved conceptualization of data uses to ensure the parameters measured are the most appropriate), and 4) maximizing the use of flow proportional autosamplers where loading data is a key end use.

At the most basic level, a clear understanding of the data end uses must be developed so that the monitoring program can be designed to collect the appropriate information. For monitoring programs that have been in existence for 10-20 years, the goals and objectives of the program may have changed, yet there are no written updates or modifications stating these revised goals and objectives of the program. Several of the existing project-specific Monitoring Plans reviewed at the start of this effort presented goals and objectives that were outdated or slightly misstated. Interviews with District staff who were end users of the data often had different views of project goals/objectives than those stated in the Monitoring Plans. The District needs to ensure continuity between the field operations staff members who draft the plans and collect the data, as well as the final data end users, so that all involved have a clear idea of why the data are being collected. Data end uses should be explicitly stated in monitoring program documentation along with the appropriate statistical treatment of the data. Rationale for the collection of each parameter, be it a mandate or how the parameter is used to support the evaluation of other parameters, should also be considered.

A second observation/recommendation that became apparent during the optimizations of several of the projects was that the use of concentration data may not be appropriate when loading information is required to meet the goals/objectives of the program. The concentration data exhibited high degrees of variability making the results of any optimization procedure questionable. One way to address this variability may be to standardize concentration data. One data standardization method that should be considered, particularly for evaluating loads, is to consider flow. If an end use is to look at loading, then the optimization must take into account flow measurements so that loads can be calculated. For projects requiring calculation of loads, the District needs to explicitly state how the loads will be calculated and how flow measurements will be attained (modeled or measured). Any considerations that need to be taken into account using flow measurements need to be stated. Maximizing the use of flow proportional autosamplers should also be considered in these instances, as well as increasing autosampler replicates to ensure representative data.

To ensure optimization approaches are adequate, it is necessary that the District define the necessary level of change that must be detected by the monitoring data from any specific program, as well as the period of time over which they need to be able to detect those changes. This needs to occur for every parameter for each monitoring project. Because it was not stated in any of the documents made available for the effort, this optimization effort considered a target change of 20% over a 5-year time period. Depending on the goals and objectives of any given monitoring program, this may or may not be sufficient. For example, if the projects optimized under this effort need to be able to detect changes in TPO4 of 5% annually, then most of the current designs are not sufficient. Identifying this information is key to any optimization effort and is critical to determining an appropriate spatial and temporal design.

As indicated previously, District scientists mentioned that several parameters could be removed from all projects. These included Chlorophyll b, Chlorophyll c, carotenoids, and ammonia from autosamplers. This optimization effort evaluated only five parameters from any given project. These parameters were selected based on discussions with District scientists, best professional judgment, and general understanding of the mission of the District. It is recommended that the District consider conducting optimization efforts on all parameters within a specific project, if those parameters are critical.

4.3 Costs/Benefits of Recommendations

The implementation of many of these recommendations may or may not result in cost savings for the District. The costs associated with the monitoring programs can be broken into three categories, field costs, laboratory costs, and data loading/maintenance costs. Depending on the specific spatial extent of some of the monitoring programs, as well as the logistics for sampling, the elimination of one or two sampling stations or a couple of parameters for a given program may not provide large savings. In contrast, removal of an entire program (i.e., RAIN) will provide a savings, but elimination of one station in an area where the stations are in close proximity will only incrementally lower laboratory and data management costs. In field sampling, preparation for the activity and getting the appropriate equipment/boats, etc., as well as the total amount of time required to complete the activities, generate the largest costs. Deploying staff in the field lowers cost only if the field crew can be downsized or 1-2 days of sampling time is eliminated. Since the sampling across a set of stations in a given day requires the same equipment, no real savings is associated with equipment investments and maintenance.

Likewise, eliminating 1-2 parameters may or may not result in considerable savings. Once in the field, collection of an additional small sample volume or conducting a field filtration does not greatly increase sampling effort. If collection of a particular parameter is very labor intensive and the data use marginal, elimination of the parameter may save some time which would equate to labor costs. Depending on the parameter, the laboratory analysis for reduction of a few samples may not provide substantial cost savings due to equipment set up and maintenance costs. Only if the reduction is large across the entire set of projects are substantive savings expected. Moreover, for methods that output several parameters, there may be little cost savings if an automatically generated parameter is eliminated. If a method for one specific parameter is costly and time intensive however, if that parameter is not critical, perhaps eliminating it or reducing the frequency with which it is collected and measured could provide some savings.

The often overlooked costs associated with monitoring programs are associated with quality assurance/quality control and the maintenance of the data. Once data are collected and analyzed, the data will still need to undergo quality checks, validation, loading into a database, and maintenance of that database. The District should also consider the cost savings associated with automating these types of activities and instituting a periodic audit of the data once it is loaded into the database. These innovations can translate to cost savings by preventing the types of database reconciliations experienced during this project. Once these types of activities have been maximally automated, there should be additional cost savings.

One cost saving practice the District may want expand upon is to further enhance the forming of partnerships with other agencies, many of whom have similar needs for data and information. Many of the current monitoring stations will be used for the RECOVER Monitoring and Assessment Plan. Perhaps, some of the funding for these programs could be covered by RECOVER funds. Additionally, for locations within the National Park boundaries, partnerships with the Department of Interior may help reduce costs. Agencies that require District data for activities such as nutrient criteria development and TMDL development may also provide an avenue for cost-sharing.

5. SUMMARY OF REVIEW COMMENTS REGARDING PREVIOUSLY CONDUCTED OPTIMIZATIONS

Five documents pertaining to water quality optimization efforts performed by District staff or by third parties under contract to the District were critically reviewed. These are:

1. Lake Okeechobee Northern Watershed Micro-Basin Sampling Network Optimization Process, February 20, 2004 by Paul D. Robillard, Ph.D.
2. WOD Monitoring Issues, No date or author listed
3. Network Optimization Questionnaire, February 2003, Bahram Charkhian
4. South Florida Water Quality Monitoring Network, No date or author listed
5. South Florida Coastal Water Quality Monitoring Network: Summary of Proposed Network Optimization, No date or author listed

Documents 1 and 2 pertain to the area north of Lake Okeechobee Northern Watershed Micro-Basin Sampling and were reviewed together. Documents 3 through 5 relate to the South Florida Coastal Monitoring Network and were reviewed together.

The following criteria were used to guide the review of these documents.

1. Identification of mandates or other monitoring objectives that motivate the monitoring project
2. Identification of current and future uses of the monitoring data by the District
3. Identification of relevant data from other monitoring projects
4. Identification of statistical analysis procedures used and acceptable levels of error
5. Selection and implementation of optimization methodologies
6. Soundness of optimization recommendations

The following sections provide a brief description of the reviews and summarize the comments relative to the criteria above:

5.1 Lake Okeechobee Northern Watershed Micro-Basin Sampling Network Optimization

Review results provided below are in the form of general comments followed by specific comments addressing each of the review criteria.

General Comments

1. Mandates and monitoring objectives associated with the subject monitoring network are very well described, as are ways in which the District will use the monitoring data. This information forms an excellent base from which to perform an optimization.

2. The optimization process recommended in this report is very general in nature. It includes some reasonable concepts for relative optimization by sampling more frequently and with increased geographical coverage in areas that exhibit higher levels of environmental impact. However, the practical and logistical issues associated with the proposed approach are not addressed, leaving the reader to wonder whether or not the approach can be implemented in practice.

3. The Kruskal-Wallis-based methodology proposed for ranking basins, sub-basins and micro-basins seems reasonable as a ranking methodology. However, the associated tests of statistical significance will only be valid if the water quality data from one station are independent of the data from another station and if there is no serial autocorrelation in the water quality data.

4. The report implies that a Critical Monitoring Network can be implemented that will provide water quality monitoring data to support trend detection, assessment of the effectiveness of implemented best management practices, identify high source micro-basins, and ultimately identify high source areas. However, there is no discussion of how to reconcile the competing objectives associated with these various goals nor is there any discussion of the monitoring resources that will be required to simultaneously achieve these different goals. The resources required could be prohibitively large.

5. Even though no optimization methods were implemented, statements are made about having sufficient data to begin the ranking process and establish phosphorus concentration trends. The bases for these claims are not conveyed in the report.

Identification of Mandates or Other Monitoring Objectives That Motivate the Monitoring Project

An excellent job is done of describing the mandates and monitoring objectives associated with the subject monitoring network.

Identification of Current and Future Uses Of the Monitoring Data by the District

The different ways the District will need to use the network data are covered very well in the report, at least implicitly if not explicitly.

Identification of Relevant Data from Other Monitoring Projects

No data from other monitoring projects are addressed in this report. While considerable mention is made of other types of data that will be required to properly analyze the network monitoring data, there is no recognition of the other, extensive monitoring efforts being conducted within the watershed . The list of other data required is good and appears to be somewhat comprehensive, but the methods by which the aggregate data will be analyzed and the implications with respect to optimization of the network are not addressed.

Identification of Statistical Analysis Procedures Used and Acceptable Levels of Error

The Kruskal-Wallis-based ranking procedure is well-described in terms of ranking entities. However, no detail is provided on the decision process for taking monitoring resources away from low-ranking entities and giving those resources to high-ranking entities. How big do the differences need to be to begin shifting resources? How far do you go in shifting resources from one set of entities to the other?

It is proposed in the document that entities be ranked based on water quality parameter levels as well as relative variability in water quality parameter levels. However, there is no guidance on what to do if the two methods lead to different conclusions regarding the shifting of monitoring resources.

Selection and Implementation of Optimization Methodologies

The optimization methodologies proposed seem reasonable from a theoretical perspective, but leave the reader wondering whether they could be easily implemented in practice. In particular, it seems that considerable resources might be required to simultaneously monitor for water quality trends, effectiveness of best management practices, and isolation of high source areas. Some discussion of the likely resources required would lend some credibility to the proposed methods. No optimization methods were actually implemented in the document reviewed[1].

[1] While no optimization methods were actually implemented in the document reviewed, the District has indicted that several of the recommendations have been implemented (Patricia Burke, personal communication, December 2005). The efficacy of the remainder of the recommendations should be carefully reviewed by the District prior to attempting their implementation.

Soundness of Optimization Recommendations

A statement is made that 10 samples is sufficient to begin the micro-basin ranking process. The basis for this statement is not conveyed in the report. It is implied that the proposed frequency of sampling is sufficient to establish phosphorous concentration trends over a 1-2 year period. The basis for this statement is not conveyed in the report. If the phosphorous concentration data exhibit serial autocorrelation, it seems unlikely that trends could be established in such a short period of time. The statistical analysis of data collected from project stations (under review in the current optimization effort of several District WQ Monitoring projects) that are closely related to those in the document, exhibited a high level of auto correlation.

5.2 South Florida Water Quality Monitoring Network

The documents reviewed for this project include a November 14, 2003 draft Network Optimization Questionnaire with supporting attachments plus a separate document prepared by FIU that contains supporting statistical results and maps (this was a primary document used to develop the questionnaire) and a summary power point presentation on the findings. The documents were reviewed with respect to the following criteria:

- Identification of mandates or other monitoring objectives that motivate the monitoring project
- Identification of current and future uses of the monitoring data by the District
- Identification of relevant data from other monitoring projects
- Identification of statistical analysis procedures used and acceptable levels of error
- Selection and implementation of optimization methodologies
- Soundness of optimization recommendations.

The review comments apply primarily to the draft Network Optimization Questionnaire as the other documents are inclusive in the draft memo. Review results are provided below in the form of general comments followed by specific comments addressing each of the review criteria.

General Comments

The stated purpose of the questionnaire was to solicit feedback from various state and federal agencies on District proposed monitoring reductions to a series of coastal monitoring program defined by geographic area. The questionnaire adequately defines its purpose and goals and conveys the statistical analysis performed (see below for more information). However, the goals of the individual projects optimized are not clearly conveyed in the document.

The analytical approach applied to develop recommendations used both professional judgment and quantitative statistical analysis. Profession judgment was used to define the principal geographic areas and sub areas of coastal south Florida. The first level of station review was guided by a set of questions that pertain to the rationale for inclusion of the stations in the monitoring programs. The document indicates that expert opinion, similarity with regards to salinity, and geographic information were used for the initial grouping of stations into geographic regions and subgroups within a geographic area, although it is unclear what specific criteria were used. The presentation alludes to, but does not specify, quantitative criteria for evaluating the importance of each station to the District's mission, environmental relevance, or types of environmental variability, gradients, and trends that are critical for decision making. This, along with potential inconsistencies in the actualization of the proposed recommendations for monitoring reductions as considered below, leave the reader with a sense of arbitrariness concerning the recommendations listed[2].

[2] Note: This optimization was performed in response to District management requests to evaluate where monitoring costs could be reduced and was conducted under a tight schedule which drove the study to a station redundancy analysis only. The fact that the document incorporates both technical and management perspectives may explain the apparent inconsistencies found by this technical review.

The level of detail and supporting tables and figures for the specific geographic areas, as well as recommendations for discontinuing stations are reasonably presented. However, the document does not address the role of spatial proximity of the stations nor autocorrelation on the power to detect trends. This is in part due to the focus of the document on station removal rather than optimization of the monitoring program and also on not having a clearly defined set of monitoring objectives, decision rules (problem statement), and questions to drive the design. For instance, if the objective of the monitoring is on trend detection, the impact of removing stations on the ability to detect trends is not discussed. This lack of specificity leads to statements that have little justification of intent. For example, on page 4 there is a statement that says the "*overall array of stations is reasonable for this complex marine environment*" and justifies retention of all stations based on the interest of other agencies. It may be more appropriate to define the changes that are expected to occur in this environment as restoration proceeds, and test the data with and without redundant stations for the ability of each design to detect long term trends. The stated objective of the overall monitoring and reason other agencies are interested in the data should be more clearly defined.

Moreover, justification of many statements in the tables could have been better supported by discussions in the text. For example, justification for the statement "*The number of stations monitoring in this area is far more than is needed to assess general water quality"* on page 5 is not provided in the document. Other statements and tables throughout the document use the term "representative of overall bay quality" yet quantitative information to enable adequate judgment of the representativeness is not presented in the report. Moreover, there is inconsistency in the information provided in the tables, figures, and text. For example, in the Ten Thousand Island discussion, Stations 57 and 58 are shown to be fully redundant by the criteria set forth in the document, yet both are retained as stations to monitor. In many other cases, this redundancy was sufficient to recommend removal of at least one of the stations. The rationale is not clear and could reflect other input not documented in the report. In contrast, other text in the document suggest that one station will be able to represent an entire subregion or that station transects that seem to be set up to sample gradients are reduced to one station without documentation as to why. There are similar statements in the Mangrove Bay discussion which states upstream stations should be eliminated on the basis of no established need. Statements in the Ten Thousand Island discussion state that "*Stations associated with freshwater discharges at the coast are* <u>*obviously*</u> (emphasis added) *most relevant to evaluating upstream, inland discharges in water quality and gulf Coast loading from such changes*" and "*the purposes of dispersing stations throughout mangrove islands and the nearshore areas…are less clear …*" which leads to a conclusion that these stations are *not likely to detect* <u>*any*</u> *(*emphasis added*) signal from freshwater sources due to the confounding marine influence …*"[3].

Other statements in the document indicate that status and trends detection objectives can be met with the recommended stations (i.e., without the redundant stations), yet no trend statistical analysis are presented to demonstrate the adequacy of the chosen stations to detect trends either at the station or subsystem level. Similarly, the rationale for excluding stations in BISC is not clearly stated. The recommendations may be better served by employing a sophisticated analysis such as that of Caccia and Boyer (2005), along with a clear definition of objectives and decisions. For example, the statement that 14 sites in Biscayne Bay are adequate to capture status and trends should be quantitatively supported. Known gradients in the system and areas that experience high variability should be accounted for quantitatively.

As written, the network questionnaire document has several poorly supported recommendations, especially with respect to the ability to conduct status and trends monitoring. However, it is clear that the District is seeking input from other agencies and potential funding partners regarding the

[3] Note that the optimization investigation conducted for this region identified that these stations will be critical in evaluation the changes and influence of upstream District projects, such as the Picayune Strand hydrological restoration.

recommendations which is an appropriate activity given the overall importance of coastal monitoring and the anticipated changes resulting from restorations activities in south Florida.

This document could be improved by including more detail on specific program objectives, the questions that drive the monitoring, and the rationale that drove the original network design. This background is essential to enable the reader to make informed judgments on the suggested changes. While some of this background information is in various places in the document, it is difficult to pick out and assumes the reader knows what the authors know. Thus, the document would benefit by consolidating this information as a theme early and succinctly in the introduction.

Statistical approach

The objective of the statistical analysis was to examine select water quality information at the stations level within small geographic subunits, rather than system level. The statistical procedures employed were designed to identify stations within geographical groups that produce similar water quality monitoring data over time. The 177 stations investigated were divided first into five geographic segments and then each geographic segment was divided into 1-5 groups again on a geographical basis. For this analysis, four key water quality parameters were chosen for station by station comparison. All pairs of stations within a group were compared first based on water quality parameter trends and then on average parameter values for salinity, chlorophyll a, total phosphorus, and total nitrogen. A station pair received an integer redundancy score in the range 0-4 indicating the number of water quality parameters for which differences between the stations were not statistically significant. Adjustments were made to control the experiment wise error rate across all the statistical tests performed. Those station pairs that were significantly different were characterized as statistically essential for the program and retained. Station pairs with all four parameters not statistically different were considered redundant and listed as candidate stations for removal from the monitoring program.

The statistical procedures employed represent a practical method for characterizing the redundancy of pairs of stations, provided that the list of parameters employed covers the range of important information content in the monitoring data. Reporting out the redundancy scores for all pair of stations in a group is an effective way of concisely characterizing the level of redundancy present in the group. However, the information provided in the report, does enable the reader to understand the specific statistical procedures used to determine whether trends and average values differed for a pair of stations. If linear regression and t-test models were used that assume no serial autocorrelation and further assume that the observations from one stations are independent of the observations from the second station, these assumptions would have to be validated to corroborate the redundancy scores.

No specific District uses of the monitoring network data uses were identified that can be translated into quantitative data quality objectives. It follows that there was no attempt to assess whether or not the current network monitoring plan was sufficient to support any specific, quantitative monitoring objectives.

I**dentification of mandates or other monitoring objectives that motivate the monitoring project**

The primary objective of the document is to solicit input from other agencies that use the District's data and conduct monitoring, presumably under mandates relevant to their mission, and to facilitate partnership and fiscal support from other agencies regarding the monitoring program. The document conveys two general purposes of the monitoring network, The first is to document status and trends of the coastal areas and the second to demonstrate progress towards protecting and restoring marine resources in South Florida (from only a water quality perspective). Potential data uses and data users are included at a general level in the project specific subsections and a general description of the types of decisions that are expected to be made from the data are noted. Environmental programs in the region are also briefly described but not linked specifically to each geographic area and the current or planned restoration projects. This implies that there are no mandated permits, legislation, or agreements that require this monitoring. Overall, the document would benefit from more clearly stating the justification of the District's specific data quality objectives either at the program level or for each project.

Identification of current and future uses of the monitoring data by the District

The document is designed to solicit input on the current and future use of the coastal monitoring data. Until this information is compiled and specific monitoring objectives and decision criteria developed, revisions to the program are problematic. There are currently several high priority District restoration projects, within the region, that would utilize this data to assess performance.

Identification of relevant data from other monitoring projects

Other monitoring program data is not included in the analysis but recognized as important to evaluate as part of the optimization, especially in Biscayne Bay where two major monitoring programs are in progress.

Identification of statistical analysis procedures used and acceptable levels of error

The statistical method used to evaluate station redundancy is reasonably described but lacks the specific information from which to judge whether trends and average values differed for a pair of stations. Also, serial autocorrelation is not evaluated and it is assumed that the observations from one station are independent of the observations from the second station. These factors should be validated to validate the redundancy scores.

Selection and implementation of optimization methodologies

The procedures employed represent a practical method for characterizing the redundancy of pairs of stations, provided that the list of parameters employed covers the range of important information content in the monitoring data. Reporting out the redundancy scores for all pair of stations in a group is an affective way of concisely characterizing the level of redundancy present in the group.

Soundness of optimization recommendations

Overall, the document is a good start towards optimization and appropriately solicits input from other stakeholders. However, it needs to be enhanced in area of monitoring objectives, questions, and decisions (as is being asked of potential end data users and partners), better conceptualization of how each system works from physical and water quality perspectives (to enable better considerations of interactions), role of gradients in the regions, and statistical analyses for the power to detect trends based on the current and proposed revisions at subsystem level rather than individual station level. Discussions on the questions posed of the external agencies (e.g., what are the Everglades National Park monitoring objectives for Whitewater Bay) are essential to the success of the optimization. Once the comments are received the current and recommended stations and stations sets should be tested for the power to detect trends against well defined decision criteria.

6. RECOMMENDATIONS FOR A DISTRICT MONITORING EVALUATION TOOL

In the process of optimizing the District's surface water quality monitoring network, Battelle accumulated large amounts of information and become familiar with overall District surface water quality monitoring needs and goals. Battelle was asked to develop recommendations for a standard approach to evaluating future water quality monitoring requests for use by District staff. The recommendations were to consider but not be limited to the following elements of water quality monitoring networks: parameter evaluation; site selection criteria; goals and objectives as they relate to the District's mission; and mandate level considerations.

The early optimization efforts for the project demonstrated that identification and articulation of the end use of the water quality monitoring data and the reports that use the data are a critical early step in this process. Moreover, application of the key elements of the EPA DQO process (USEPA 2000) was identified as a critical activity for the optimization process. Based on these experiences, the following steps should be followed by the District for designing new monitoring programs or evaluating existing programs:

Step 1: Clearly define project objectives and goals, describe all data uses for the monitoring project, and state the management and policy decisions the data will support.

Step 2: Ensure any data used to run statistical analyses are appropriate, complete and accurate.

Step 3: Incorporate the EPA DQO stepwise approach when designing a monitoring plan or undertaking revisions to current monitoring programs,

Step 4: Define the geographic domains and whether the data for individual sites or geographic regions are to be evaluated.

Step 5: Address seasonal trends and autocorrelation to ensure statistical results are not overstating the power of the monitoring to detect trends.

Step 6: Apply the SAS power analysis procedure for trend detection. (See below for a summary description)

It is also imperative that the District understand how the various parameters inform not only the District's goals and objectives but those defined by the project. It is also critical that each parameter proposed for monitoring has a clearly defined use and purpose. These can range from a permit requirement to a parameter that supports interpretation of key permit parameters. Parameters that can not be placed into such context should not be considered for inclusion in the program.

Under this optimization project Battelle identified that one of the most common water quality monitoring objectives that motivates monitoring performed by or on behalf of the South Florida Water Management District is detection of an increasing or decreasing trend in a water quality parameter. A prominent example is the monitoring of total phosphorus (TPO4) concentrations in surface water in the Kissimmee and Okeechobee watersheds. The Lake Okeechobee Protection Plan (LOPP) calls for a 70% reduction in the TPO4 load to Lake Okeechobee by 2015 and a near-shore TPO4 concentration of less than 40 ppb (µg/L). The LOPP also specifies construction projects, management projects, and a myriad of best management practices that are designed to achieve these TPO4 goals. Over the next decade, the District will use its water quality monitoring data and statistical trend analysis procedures to assess the effectiveness of LOPP implementation toward meeting the 2015 TPO4 goals. Trends in TPO4 concentration and load will be assessed at basin, sub-basin and tributary levels.

Battelle
The Business of Innovation

A key question related to the District water quality monitoring network is whether or not the monitoring data collected will be sufficient to assess the effectiveness of projects and practices implemented to control and improve water quality and determine whether or not sufficient progress is being made toward water quality goals and objectives. One way to address this question is to perform statistical power analyses to determine the smallest water quality trends that will be detectable with high probability based on water quality data collected according to current monitoring plans. Using the resulting detectable trends, District staff will be able to determine whether the trends necessary to achieve long-term goals will be discernable from trends that fail to achieve the long-term goals.

Battelle developed a power analysis procedure and SAS program called *trend_power.sas* to facilitate performance of statistical power analyses for trend detection by District staff. The basic power analysis procedure involves the following steps:

- Fit a statistical model to the data to have a basis for generating simulated water quality parameter data to support a Monte Carlo based power analysis procedure
- Generate multiple replicate simulated water quality time series data sets
- Perform a Seasonal Kendall's Tau trend analysis procedure (Reckhow et al. 1993) for each simulated time series data set; in particular, obtain a point estimate of the slope vs. time for the log-transformed water quality parameter values
- Estimate the *annual proportion change* (APC) in water quality parameter values that is detectable with 80% power using a simple two-sided test based on the slope estimate performed at a 5% significance level

A detailed description of the proposed procedure is provided in the steps that follow.

1. Check Assumptions. The power analysis procedure proposed here applies to time series data that, once log-transformed, follows a linear trend over time. Visually check to make certain that the log-transformed water quality data exhibit homogeneous variability over time about a simple linear long-term time trend model and that there are no overly influential outliers. Remove overly influential outliers. If any model assumptions are violated, possible options are:

    1. In Step 2, fit a mixed model that contains a more complicated fixed effects component (requires software modification), and/or

    2. Select a subset of the data that satisfies the model assumptions

2. Fit Mixed Model to the Water Quality Time Series Data. Using SAS PROC MIXED, fit a mixed model to the natural log-transformed water quality data to produce a set of model parameters for use in simulating data.

    The mixed model fitted to the data is specified as follows.

    $$Y_t = \alpha + \beta(t\text{-}t_0) + S_t + \varepsilon_1 + \varepsilon_2$$

    where

    Y_t = natural log-transformed water quality measurement at time t

α = average seasonally-adjusted water quality measurement value at time t_0

β = average change in water quality measurement per unit time;

t = time of sample collection;

t_0 = reference time point to give relevance to the α parameter

S_t = seasonal effect at time t that repeats itself on an annual cycle and averages to zero;

ε_1 = random error term (with mean zero and standard deviation σ_1) associated with temporal variability in true water quality measurement values; and

ε_2 = random error term (with mean zero and standard deviation σ_2) associated with sampling and chemical analysis variability.

The ε_2 error terms are assumed to be stochastically independent from sample to sample whereas the correlation between the ε_1 error terms at times t_1 and t_2 is assumed to be equal to

$$\rho^{|t_2 - t_1|}.$$

The model is fitted to log-transformed water quality parameter measurements instead of the measurements themselves for two reasons. First, and most important, many of the District's water quality parameters are concentrations of compounds or elements in water and it is our experience that environmental concentrations tend to be more log-normally distributed than normally distributed. Modeling log-transformed parameter measurements, therefore, generally increases the validity of models with normally distributed error terms such as the model used to simulate data as part of the Monte Carlo based power analysis procedure proposed here. Second, modeling log-transformed parameter measurements allows detectable trends to be stated in terms of percentage changes rather than absolute changes in parameter values, causing the detectable change results to be more easily interpretable.

For some time series, the full mixed model specified above cannot be fitted to the data because of convergence problems associated with SAS PROC MIXED. In this case it is recommended that reduced model that excludes the ε_2 term be fitted to the data.

The average annual proportion change (APC) in water quality parameter value can be expressed as a function of the slope parameter β:

$$APC = \exp(\beta) - 1$$

3. Simulate Monitoring Data for a Specified Monitoring Design. Simulate monitoring data according to a specified monitoring design using the mixed model parameters from Step 2 except replace the *annual percentage change model parameter* with the value 0 (equivalently β=0). The key parameters used to specify the monitoring design are:

    - The number of years over which data should be generated
    - The number of seasons per year for which data is generated
    - The probability that a sample will not be obtained at a specified sampling time

The latter parameter value allows one to incorporate known frequencies of “No Bottle Sample” occurrences (and other causes of missing data) into the power analysis.

4. Estimate the *Annual Proportion Change* Detectable with 80% Power Employing a Specified Monitoring Design. For the simulated data sets generated in Step 3, perform a statistical test for trend based on the Seasonal Kendall’s Tau slope estimator and an assumed normal distribution. Employ a 5% significance level when performing the test. Estimate the *slope parameter value* detectable with 80% power. Calculate the *annual proportion change* (APC) detectable with 80% power based on the formula

$$\text{detectable\_APC} = \exp(\text{detectable\_slope}) - 1$$

Alternatively, the statistical test for trend could have been based on a parametric model for water quality parameter values such as the model proposed in Step 2 and used to generate data in Step 3. However, such models have the potential to be unduly influenced by departures from distributional assumptions, particularly extreme outliers in the monitoring data sets. To avoid these potential problems, the statistical test for trend was based on the nonparametric seasonal Kendall’s Tau procedure and the accompanying robust median slope estimator.

A detailed description of the procedure and the code was submitted under separate cover to the District and will be made available on the South Florida Water Management District’s Environmental Resource Assessment Department’s internal web link.

7. REFERENCES

Caccia, V.G. and Boyer, K.N. 2005. Spatial patterning of water quality in Biscayne Bay, Florida as a function of land use and water management. *Mar. Poll. Bull*. 50:1416-1429.

Battelle. 2005. Comprehensive Water Quality Monitoring Network Optimization and Quality Control Optimization. Work Plan for South Florida Water Management District, Environmental Resource Assessment Department Work Order Number C-15968-WO04-11, April 2005.

Hatcher, L. and E. J. Stepanski. 1994. *A Step by Step Approach to Using the SAS System for Univariate and Multivariate Statistics*. Cary, NC. SAS Institute Inc. 552pp.

Reckhow, KH, Kepford K, and Hicks WW (1993). Methods for the Analysis of Lake Water Quality Trends. EPA 841-R-93-003.

Rust, SW. (2005). Power Analysis Procedure for Trend Detection with Accompanying SAS Software. Battelle Report to South Florida Water Management District, November 2005.

U.S. Environmental Protection Agency. 2000.. Guidance for the Data Quality Objectives Process. Report No. EPA/600/R-96/055. Washington, DC: U.S. EPA Office of Environmental Information.

Zar, J.H. 1984 *Biostatistical Analysis.* Prentice Hall. Englewood Cliffs, New Jersey. 07632

ATTACHMENT 1: SAS TREND ASSESSMENT PROTOCOL SUMMARY

POWER ANALYSIS PROCEDURE FOR TREND DETECTION WITH ACCOMPANYING SAS SOFTWARE

Steven W. Rust
January 4, 2005

1.0 INTRODUCTION

One of the most common water quality monitoring objectives that drives monitoring performed by or on behalf of the South Florida Water Management District (District) is detection of an increasing or decreasing trend in a water quality parameter. A prominent example is the monitoring of total phosphorus (TPO4) concentrations in surface water in the Kissimmee and Okeechobee watersheds. The Lake Okeechobee Protection Plan (LOPP) calls for a 70% reduction in the TPO4 load to Lake Okeechobee by 2015 and a near-shore TPO4 concentration of less than 40 ppb (µg/L). The LOPP also specifies construction projects, management projects, and a myriad of best management practices that are designed to achieve these TPO4 goals. Over the next decade, the District will use its water quality monitoring data and statistical trend analysis procedures to assess the effectiveness of LOPP implementation toward meeting the 2015 TPO4 goals. Trends in TPO4 concentration and load will be assessed at basin, sub-basin and tributary levels.

A key question related to the District water quality monitoring network is whether or not the monitoring data collected will be sufficient to assess the effectiveness of projects and practices implemented to control and improve water quality and determine whether or not sufficient progress is being made toward water quality goals and objectives. One way to address this question is to perform statistical power analyses to determine the smallest water quality trends that will be detectable with high probability based on water quality data collected according to current monitoring plans. Using the resulting detectable trends, District staff will be able to determine whether the trends necessary to achieve long-term goals will be discernable from trends that fail to achieve the long-term goals.

To facilitate performance of statistical power analyses for trend detection by District staff, Battelle has developed a power analysis procedure and written a SAS program called *trend_power.sas* (see Appendix A for a program listing) that can be used to implement the procedure. The procedure is detailed in Section 2. Documentation of *trend_power.sas* is included in Appendix B and discussed in Section 3. Finally, in section 4, an example is provided that illustrates the use of *trend_power.sas* with the data in Appendix C.

2.0 POWER ANALYSIS PROCEDURE FOR TREND DETECTION

The basic power analysis procedure proposed here involves the following steps:

- Fit a statistical model to the data in order to have a basis for generating simulated water quality parameter data to support a Monte Carlo based power analysis procedure
- Generate multiple replicate simulated water quality time series data sets
- Perform a seasonal Kendall's Tau trend analysis procedure (Reckhow et al. 1993) for each simulated time series data set; in particular, obtain a point estimate of the slope vs. time for the log-transformed water quality parameter values
- Estimate the *annual proportion change* (APC) in water quality parameter values that is detectable with 80% power using a simple two-sided test based on the slope estimate performed at a 5% significance level

A detailed description of the proposed procedure is provided in the steps that follow.

1. Check Assumptions. The power analysis procedure proposed here applies to time series data that, once log-transformed, follows a linear trend over time. Visually check to make certain that the log-transformed water quality data exhibit homogeneous variability over time about a simple linear long-term time trend model and that there are no overly influential outliers. Remove overly influential outliers. If any model assumptions are violated, possible options are:

    1. In Step 2, fit a mixed model that contains a more complicated fixed effects component (requires software modification), and/or

    2. Select a subset of the data that satisfies the model assumptions

2. Fit Mixed Model to the Water Quality Time Series Data. Using SAS PROC MIXED, fit a mixed model to the natural log-transformed water quality data to produce a set of model parameters for use in simulating data.

    The mixed model fitted to the data is specified as follows.

    $$Y_t = \alpha + \beta(t-t_0) + S_t + \varepsilon_1 + \varepsilon_2$$

    where

    Y_t = natural log-transformed water quality measurement at time t

    α = average seasonally-adjusted water quality measurement value at time t_0

    β = average change in water quality measurement per unit time;

    t = time of sample collection;

    t_0 = reference time point to give relevance to the α parameter

S_t = seasonal effect at time t that repeats itself on an annual cycle and averages to zero;

ε_1 = random error term (with mean zero and standard deviation σ_1) associated with temporal variability in true water quality measurement values; and

ε_2 = random error term (with mean zero and standard deviation σ_2) associated with sampling and chemical analysis variability.

The ε_2 error terms are assumed to be stochastically independent from sample to sample whereas the correlation between the ε_1 error terms at times t_1 and t_2 is assumed to be equal to

$$\rho^{|t_2 - t_1|}.$$

The model is fitted to log-transformed water quality parameter measurements instead of the measurements themselves for two reasons. First, and most important, many of the District's water quality parameters are concentrations of compounds or elements in water and it is our experience that environmental concentrations tend to be more log-normally distributed than normally distributed. Modeling log-transformed parameter measurements, therefore, generally increases the validity of models with normally distributed error terms such as the model used to simulate data as part of the Monte Carlo based power analysis procedure proposed here. Second, modeling log-transformed parameter measurements allows detectable trends to be stated in terms of percentage changes rather than absolute changes in parameter values, causing the detectable change results to be more easily interpretable.

For some time series, the full mixed model specified above cannot be fitted to the data because of convergence problems associated with SAS PROC MIXED. In this case it is recommended that reduced model that excludes the ε_2 term be fitted to the data.

The average annual proportion change (APC) in water quality parameter value can be expressed as a function of the slope parameter β:

$$APC = \exp(\beta) - 1$$

3. Simulate Monitoring Data for a Specified Monitoring Design. Simulate monitoring data according to a specified monitoring design using the mixed model parameters from Step 2 except replace the *annual percentage change model parameter* with the value 0 (equivalently β=0). The key parameters used to specify the monitoring design are:

   - The number of years over which data should be generated
   - The number of seasons per year for which data is generated
   - The probability that a sample will not be obtained at a specified sampling time

   The latter parameter value allows one to incorporate known frequencies of "No Bottle Sample" occurrences (and other causes of missing data) into the power analysis.

4. Estimate the *Annual Proportion Change* Detectable with 80% Power Employing a Specified Monitoring Design. For the simulated data sets generated in Step 3, perform a statistical test for trend based on the seasonal Kendall's Tau slope estimator and an assumed normal distribution. Employ a 5% significance level when performing the test. Estimate the *slope parameter value*

detectable with 80% power. Calculate the *annual proportion change* (APC) detectable with 80% power based on the formula

detectable_APC = exp(detectable_slope) - 1.

Alternatively, the statistical test for trend could have been based on a parametric model for water quality parameter values such as the model proposed in Step 2 and used to generate data in Step 3. However, such models have the potential to be unduly influenced by departures from distributional assumptions, particularly extreme outliers in the monitoring data sets. To avoid these potential problems, the statistical test for trend was based on the nonparametric seasonal Kendall's Tau procedure and the accompanying robust median slope estimator.

3.0 SOFTWARE DOCUMENTATION

Battelle has developed a SAS program called *trend_power.sas* that implements the power analysis procedure specified in Section 2. See Appendix A for a program listing. The program is comprised of seven SAS macros followed by a short "main program". Most users will utilize the program according to the following procedural steps.

A. Create a SAS data set containing the time series data that forms the basis for the power analyses to be performed

B. Specify one or more calls to the *run_one_sim* macro in the "main program" section

C. Examine the printed output or perform additional analyses on the permanent SAS data set (*wqdata.trend_power_results*) that contains the power analysis results

In Appendix B, detailed documentation is provided for Steps A-C above and each of the seven SAS macros and the main program section included in the *trend_power.sas* program.

In addition to implementing the power analysis procedure specified in Section 2, the *run_one_sim* macro can be employed to obtain information about the true significance levels and power values associated with three variations of the seasonal Kendall's Tau test for trend. The macro performs the following tests for trend:

1. Seasonal Kendall's Tau test without correction for serial autocorrelation
2. Seasonal Kendall's Tau test with correction for serial autocorrelation only when a screening test indicates the presence of statistically significant autocorrelation
3. Seasonal Kendall's Tau test with correction for serial autocorrelation always applied

If the *run_one_sim* macro is run with an APC value of zero, the true significance levels of the above seasonal Kendall's Tau tests are estimated by the proportion of simulated data sets for which the test for trend declares a statistically significant trend. If the *run_one_sim* macro is run with an APC value different from zero, the true power values of the above seasonal Kendall's Tau tests for detecting the specified APC value are estimated by the proportion of simulated data sets for which the test for trend declares a statistically significant trend.

NOTE: THE CURRENT VERSION OF THE *TREND_POWER.SAS* PROGRAM DOES NOT EXECUTE TESTS 2 AND 3 SPECIFIED ABOVE IF THE INPUT PARAMETER SPECIFYING THE PROBABILITY OF A MISSING SAMPLE IS GREATER THAN ZERO.

4.0 A BRIEF EXAMPLE

As supplied with this report, the *trend_power.sas* program includes three calls to the *run_one_sim* macro that specify power analyses based on the example data set listed in Appendix C and supplied with this report as a file named *example.sas7bdat*. When executed, the program should produce the output in Figure 1 and six graphs. The first two graphs are illustrated as Figures 2 and 3. The columns in the output table are defined at the end of the documentation in Appendix B.

5.0 REFERENCES

Reckhow KH, Kepford K, and Hicks WW (1993). Methods for the Analysis of Lake Water Quality Trends. EPA 841-R-93-003.

Simulation Results

Obs	label	num_seasons_actual	sigma1sq	rho	sigma2sq	osl_resid_norm	num_years_sim	num_seasons_sim	num_reps	apc	power_autocorr	power_without	power_comb	power_with	osl_slope_norm	stderr	power_stderr	detectable_apc	pmiss
1	KREA 01 TP04	24	0.41967	0.53755	0.15923	0.40018	5	24	1000	0	0.249	0.138	.	.	0.79839	0.093297	0.05	0.29872	0.22
2	KREA 04 TP04	12	0.03604	0.69925	0.24725	0.00000	5	24	1000	0	0.005	0.067	.	.	0.60238	0.077044	0.05	0.24091	0.49
3	KREA 04 TP04	12	0.03604	0.69925	0.24725	0.00000	5	24	1000	0	0.057	0.128	0.124	.009	0.55431	0.053297	0.05	0.16104	0.00

Figure 1. Output Produced by *trend_power.sas* for Example Data Listed in Appendix C

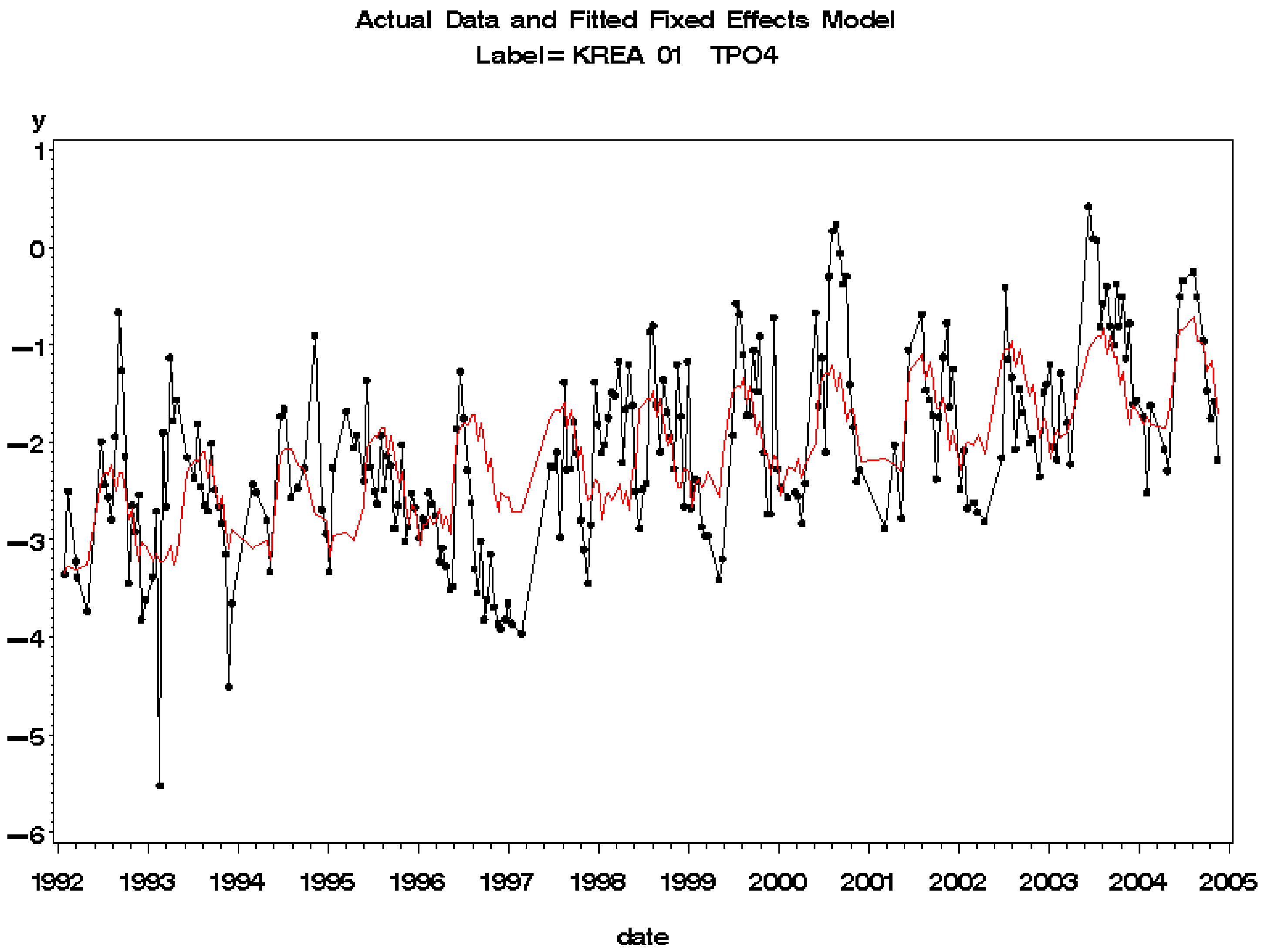


Figure 2. First Plot Produced *trend_power.sas* for Example Data Listed in Appendix C

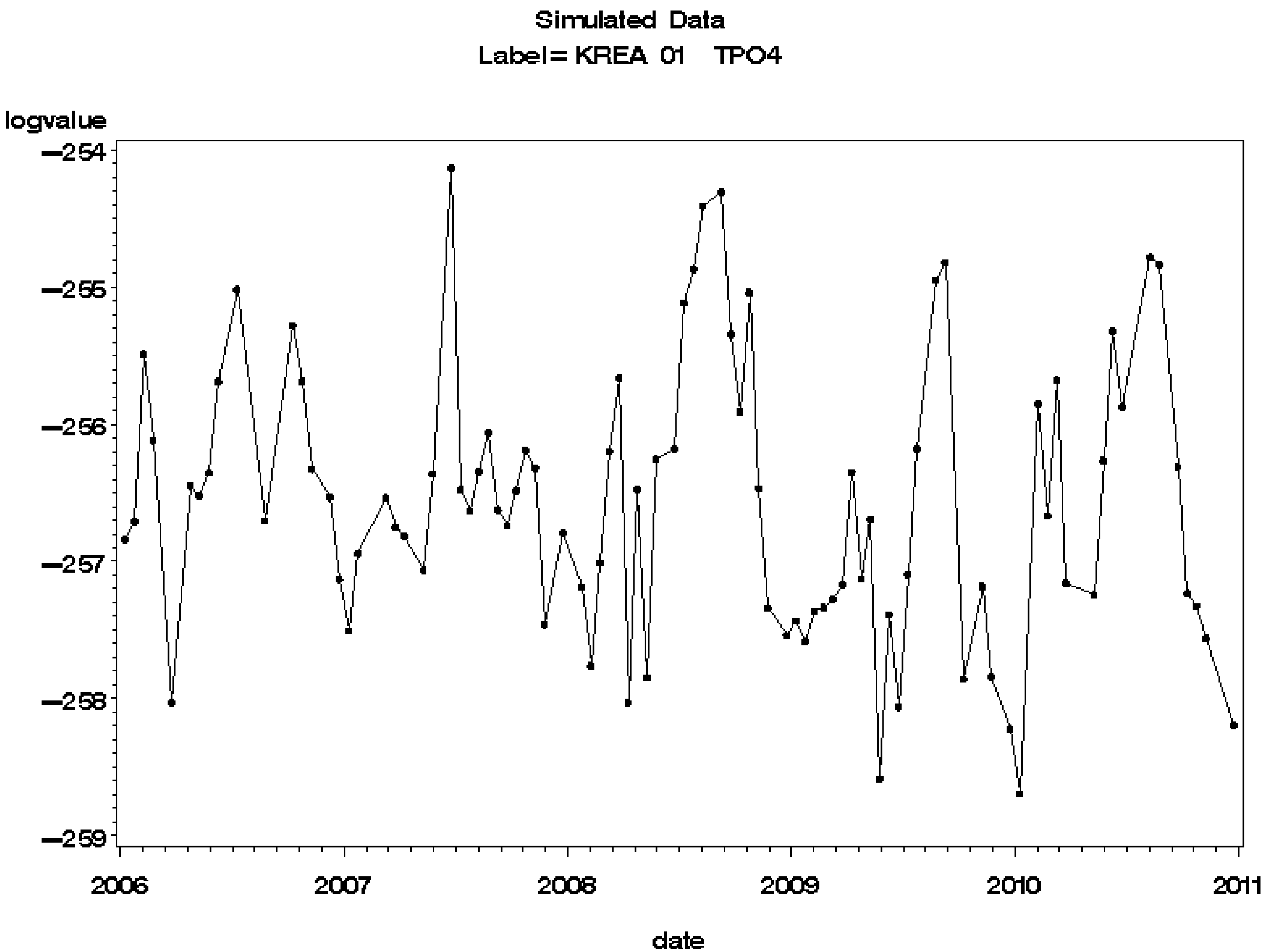


Figure 3. Second Plot Produced *trend_power.sas* for Example Data Listed in Appendix C

APPENDIX A to Attachment 1

SAS CODE LISTING FOR *TREND_POWER.SAS*

```
*******************************************************************************;
** CREATE MACRO TO FIT A MIXED MODEL WITH A LOCAL NUGGET IF IT WILL CONVERGE, **;
** AND A MIXED MODEL WITH NO LOCAL NUGGET OTHERWISE                          **;
*******************************************************************************;

%macro mixed_model(label=,indata=,parameter=,condition=,num_seasons=,outparms=,outdata=);

    ** SUBSET THE DATA TO THE TIME SERIES OF INTEREST **;

    data ts;
        set &indata;
        keep label y season date date_months;
        length label $25;
        label="&label";
        y=&parameter;
        if (&condition);
    run;

    ** FIT A TIME SERIES MODEL WITH A SEASONALITY EFFECT AND A LOCAL NUGGET **;

    proc mixed data=ts covtest;
        by label;
        class season;
        model y = season date / solution ddfm=kenwardroger outpm=&outdata residual;
        repeated / subject=intercept type=sp(pow)(date_months) local;
        ods output CovParms=covparms SolutionF=solutionf;
        title "Mixed Model With Local Nugget";
    run;

    ** IF FIRST MODEL DID NOT CONVERGE, FIT A TIME SERIES MODEL WITH A **;
    ** SEASONALITY EFFECT AND NO LOCAL NUGGET                          **;

    %let mod1_conv=%sysfunc(exist(solutionf));
    %if &mod1_conv=0 %then %do;

        proc mixed data=ts covtest;
            by label;
            class season;
            model y = season date / solution ddfm=kenwardroger outpm=&outdata residual;
            repeated / subject=intercept type=sp(pow)(date_months);
            ods output CovParms=covparms SolutionF=solutionf;
            title "Mixed Model With NO Local Nugget";
        run;

    %end;

    ** TEST THE RESIDUALS FOR NORMALITY **;

    proc univariate data=&outdata noprint;
        var resid;
        output out=osl_resid_norm probn=osl_resid_norm;
    run;

    ** REFORMAT MIXED MODEL OUTPUT **;

    data covparms_2;
        set covparms;
        retain sigma1sq osl_sigma1 rho osl_rho;
        keep sigma1sq osl_sigma1 rho osl_rho sigma2sq osl_sigma2;

        if (&mod1_conv=1) then do;
            if (covparm="Variance") then do;
                sigma1sq=estimate;
                osl_sigma1=probz;
            end;
            if (covparm="SP(POW)") then do;
                rho=estimate;
                osl_rho=probz;
```

```
            end;
            if (covparm="Residual") then do;
                sigma2sq=estimate;
                osl_sigma2=probz;
                output;
            end;
        end;

        if (&mod1_conv=0) then do;
            if (covparm="SP(POW)") then do;
                rho=estimate;
                osl_rho=probz;
            end;
            if (covparm="Residual") then do;
                sigma1sq=estimate;
                osl_sigma1=probz;
                sigma2sq=0;
                osl_sigma2=.;
                output;
            end;
        end;
    run;

    data solutionf_2;
        set solutionf;
        keep a s1-s&num_seasons b stderr_b df osl_b;
        retain a s1-s&num_seasons;
        array s[&num_seasons] s1-s&num_seasons;

        if (effect="Intercept") then a=estimate;
        if (effect="season") then s[season]=estimate;

        if (effect="date") then do;
            b=estimate;
            stderr_b=stderr;
            df=df;
            osl_b=probt;
            output;
        end;
    run;

    ** MERGE VARIANCE COMPONENT, FIXED EFFECT, AND NORMALITY TEST RESULTS **;

    data &outparms;
        merge covparms_2 solutionf_2 osl_resid_norm;
        length label $25;
        label="&label";
        mod1_conv=&mod1_conv;
    run;

    ** DELETE THE SOLUTIONF DATA SET **;

    proc datasets library=work nolist;
        delete solutionf;
    run;
    quit;

    ** PLOT THE TIME SERIES DATA AND THE FITTED FIXED EFFECTS MODEL **;

    goptions reset=all device=win targetdevice=winprtc rotate=landscape ftext=swissb htext=0.5 cm
noprompt;
    symbol1 i=join v=dot h=0.5 l=1 c=black;
    symbol2 i=join v=none l=1 c=red;

    proc gplot data=&outdata;
        by label;
        plot (y pred)*date / overlay;
        title "Actual Data and Fitted Fixed Effects Model";
        label label="Label";
    run;
    quit;

    ** PRINT OUT OBSERVATIONS WITH ABSOLUTE STUDENTIZED RESIDUALS GREATER THAN 3 **;

    proc print data=&outdata;
        where (abs(studentresid) gt 3);
        by label;
        var date y;
        title "Observations with Absolute Studentized Residuals > 3";
    run;
```

```
%mend mixed_model;

****************************************************************************;
** CREATE A MACRO TO SIMULATE DATA BASED ON A SET OF MIXED MODEL PARAMETERS **;
****************************************************************************;

%macro
sim_data(label=,apc=,sim_parms=,num_seasons_actual=,num_years_sim=,num_seasons_sim=,num_obs=,num_
reps=,outdata=,pmiss=);

    ** SIMULATE MONITORING DATA **;

    data &outdata;
        set &sim_parms;
        retain e1;
        keep label rep num_years num_seasons y1-y&num_obs;
        length label $25;
        array s[&num_seasons_actual] s1-s&num_seasons_actual;
        array y[&num_years_sim,&num_seasons_sim] y1-y&num_obs;

        label="&label";
        num_years=&num_years_sim;
        num_seasons=&num_seasons_sim;

        slope=log(1+&apc);
        sigma1=sqrt(sigma1sq);
        sigma2=sqrt(sigma2sq);

        do rep=1 to &num_reps;
            e1=sigma1*rannor(0);
            do year=1 to &num_years_sim;
                do season=1 to &num_seasons_sim;
                    date = year + (season-0.5)/&num_seasons_sim;
                    date_months = 12 * date;
                    if (rho gt 0) then corr=rho**(&num_seasons_actual/&num_seasons_sim);
                    if (rho eq 0) then corr=0;
                    if (rho lt 0) then corr=-((-rho)**(&num_seasons_actual/&num_seasons_sim));
                    e1=corr*e1+sqrt(1-corr*corr)*sigma1*rannor(0);
                    e2=sigma2*rannor(0);
                    season_actual=round((&num_seasons_actual/&num_seasons_sim)*(season-0.5)+0.5);
                    fixed = a + s[season_actual] + slope*date;
                    y[year,season] = fixed + e1 + e2;
                    if (ranuni(0) lt &pmiss) then y[year,season]=.;
                end;
            end;
            output;
        end;
    run;

    ** PLOT FIRST REP **;

    data firstrep;
        set &outdata;
        keep label date logvalue;
        array y[&num_years_sim,&num_seasons_sim] y1-y&num_obs;

        if (rep=1) then do;
            do year=1 to &num_years_sim;
                do season=1 to &num_seasons_sim;
                    date=2005+year+(season-0.5)/&num_seasons_sim;
                    logvalue=y[year,season];
                    output;
                end;
            end;
        end;
    run;

    symbol1 i=join v=dot h=0.5 l=1 c=black;
    proc gplot data=firstrep;
        by label;
        plot logvalue*date;
        title "Simulated Data";
        label label="Label";
    run;
    quit;

%mend sim_data;

*******************************************************;
** CREATE A MACRO TO PERFORM A MANN-KENDALL PROCEDURE **;
*******************************************************;
```

```
%macro kendall(indata,nobs,ny,ns,dv,outdata,pmiss);

*******************************************************************************************************;
*******************************************************************************************************;
**                                                                                                   **;
** Functionality: Performs seasonal Mann-Kendall calculations for each observation (row) in          **;
**    an input data set.  Places results in an output data set.                                      **;
**                                                                                                   **;
** Arguments:                                                                                        **;
**    indata  = Input data set name                                                                  **;
**    nobs    = Number of data points for dimensioning (must be a number, not a variable name)       **;
**    ny      = Number of years for dimensioning (must be a number, not a variable name)             **;
**    ns      = Number of seasons for dimensioning (must be a number, not a variable name)           **;
**    dv      = Dependent variable                                                                   **;
**    outdata = Output data set name                                                                 **;
**                                                                                                   **;
** Input data format: The input data set must store the depenedent variable data in a numbered       **;
**    range list of variables for which numbering starts at one.  The name associated with the       **;
**    list id specified by the dv argument.                                                          **;
**                                                                                                   **;
** Output data format: The output data set contains the following four variables.                    **;
**    tau       = Proportion of paired comparisons with positive slope minus proportion of paired    **;
**                     comparisons with negative slope                                               **;
**    pwithout1 = observed significance level for H0: E(tau)=0 assuming no serial autocorrelation,   **;
**                     exactly as kendall3.exe performs the calulations                              **;
**    pwithout2 = observed significance level for H0: E(tau)=0 assuming no serial autocorrelation,   **;
**                     modified to change unusual aspect of kendall3.exe calulations                 **;
**    pwith1    = observed significance level for H0: E(tau)=0 correcting for serial autocorrelation, **;
**                     exactly as kendall3.exe performs the calulations                              **;
**    pwith2    = observed significance level for H0: E(tau)=0 correcting for serial autocorrelation, **;
**                     modified to change unusual aspect of kendall3.exe calulations                 **;
**    slope     = median of all within-season paired slope estimates                                 **;
**                                                                                                   **;
*******************************************************************************************************;
*******************************************************************************************************;

    data kendall(keep=tau pwithout1 pwith1 pwithout2 pwith2) slopes(keep=rep slope_ijk);
        set &indata;

        array &dv[&ny,&ns] y1-y&nobs;
        array r[&ny,&ns] r1-r&nobs;
        array nyk[&ns] nyk1-nyk&ns;

        ** Calculate tau, pwithout **;

        t=0;
        ncompsum=0;
        do k=1 to &ns;
            tk=0;
            ncompk=0;
            do i=1 to &ny-1;
                do j=i+1 to &ny;
                    if (&dv[j,k] ne . and &dv[i,k] ne .) then do;
                        tk=tk+sign(&dv[j,k]-&dv[i,k]);
                        ncompk=ncompk+1;
                        slope_ijk=(&dv[j,k]-&dv[i,k])/(j-i);
                        output slopes;
                    end;
                end;
            end;
            t=t+tk;
            ncompsum=ncompsum+ncompk;
        end;
        tau=t/ncompsum;

        vart_without=0;
        do k=1 to &ns;
            nyk[k]=0;
            do i=1 to &ny;
                if (&dv[i,k] ne .) then nyk[k]=nyk[k]+1;
            end;
            vart_without=vart_without+nyk[k]*(nyk[k]-1)*(2*nyk[k]+5)/18;
        end;

        absz_without1=max(0,abs(t)-1)/sqrt(vart_without);
        pwithout1=2*(1-probnorm(absz_without1));

        absz_without2=max(0,abs(t)-0.5)/sqrt(vart_without);
        pwithout2=2*(1-probnorm(absz_without2));
```

```
    ** Calculate ranks **;

    do k=1 to &ns;
        do i=1 to &ny;
            r[i,k]=(&ny+1)/2;
            do j=1 to &ny;
                if (&dv[j,k] ne . and &dv[i,k] ne .) then r[i,k]=r[i,k]+sign(&dv[i,k]-&dv[j,k])/2;
            end;
        end;
    end;

    ** Calculate variance including covariance terms **;

    vart_with=vart_without;
    do k=1 to &ns-1;
        do m=k+1 to &ns;

            ** Calculate skm term **;
            skm=0;
            do i=1 to &ny-1;
                do j=i+1 to &ny;
                    skm=skm+sign((&dv[j,k]-&dv[i,k])*(&dv[j,m]-&dv[i,m]));
                end;
            end;

            ** Calculate rank cross-product term **;

            rcpkm=0;
            do i=1 to &ny;
                rcpkm=rcpkm+r[i,k]*r[i,m];
            end;

            ** Calculate covariance and add it to variance **;

            covkm=(skm+4*rcpkm-&ny*(nyk[k]+1)*(nyk[m]+1))/3;
            vart_with=vart_with+2*covkm;
        end;
    end;

    if (vart_with gt 0) then absz_with1=max(0,abs(t)-1)/sqrt(vart_with);
        else absz_with1=0;
    pwith1=2*(1-probnorm(absz_with1));

    if (vart_with gt 0) then absz_with2=max(0,abs(t)-0.5)/sqrt(vart_with);
        else absz_with2=0;
    pwith2=2*(1-probnorm(absz_with2));

    ** Since "with" analysis is not working for missing data, **;
    ** make "with" results missing if there is missing data   **;
    if (&pmiss gt 0) then do;
        pwith1=.;
        pwith2=.;
    end;

    output kendall;
run;

** CALCULATE THE (MEDIAN) SLOPE ESTIMATE **;

proc means data=slopes noprint;
    by rep;
    var slope_ijk;
    output out=medslope(drop=_type_ _freq_) median=slope;
run;

** MERGE THE TAU AND SLOPE RESULTS **;

data &outdata;
    merge kendall medslope;
run;

%mend kendall;

*****************************************************************;
** CREATE A MACRO TO DETERMINE THE PRESENCE OF AUTOCORRELATION **;
*****************************************************************;

%macro autocorr(indata=,medslope=,num_years=,num_seasons=,num_obs=,outdata=);

    ** CREATE LONG DATA SET OF DE-TRENDED OBSERVATIONS **;
```

```
data long;
    merge &indata &medslope(keep=rep slope);
    by rep;
    keep rep year season ydt;
    array ym[&num_years,&num_seasons] y1-y&num_obs;

    do year=1 to &num_years;
        do season=1 to &num_seasons;
            date=(year-1)+(season-0.5)/&num_seasons;
            ydt=ym[year,season]-slope*date;
            output;
        end;
    end;
run;

** SUBTRACT OFF SEASONAL MEDIANS **;

proc sort data=long nodupkey;
    by rep season year;
run;

proc means data=long noprint;
    by rep season;
    var ydt;
    output out=seasonal_medians(drop=_type_ _freq_) median=seasmed;
run;

data long_2;
    merge long seasonal_medians;
    by rep season;
    keep rep year season ydtds;
    ydtds=ydt-seasmed;
run;

proc sort data=long_2 nodupkey;
    by rep year season;
run;

** CALCULATE REGRESSION COEFFICIENTS, OBSERVED SIGNIFICANCE LEVELS, **;
** AND AUTOCORRELATION INDICATOR                                     **;

data lags;
    set long_2;
    by rep;
    retain count;

    if first.rep then count=0;
    count=count+1;

    lag0=ydtds;
    lag1=lag(ydtds);
    lag2=lag2(ydtds);

    if (count=1) then do;
        lag1=.;
        lag2=.;
    end;
    if (count=2) then lag2=.;
run;

proc reg data=lags outest=regout/*(keep=rep _type_ _depvar_ lag0)*/ tableout noprint;
    by rep;
    model lag1=lag0;
    model lag2=lag0;
run;
quit;

data &outdata;
    set regout;
    keep rep b1 p1 b2 p2 autocorr;
    retain b1 p1 b2;
    if (_type_="PARMS"  and _depvar_="lag1") then b1=lag0;
    if (_type_="PVALUE" and _depvar_="lag1") then p1=lag0;
    if (_type_="PARMS"  and _depvar_="lag2") then b2=lag0;
    if (_type_="PVALUE" and _depvar_="lag2") then do;
        p2=lag0;
        if (b1 gt 0 and p1 le 0.1 and b2 gt 0 and p2 lt 0.1) then autocorr=1;
            else autocorr=0;
        output;
    end;
run;
```

```
%mend autocorr;

*****************************************************;
** CREATE A MACRO TO RUN A FULL KENDALL SIMULATION **;
*****************************************************;

%macro
sim_kendall(label=,apc=,sim_parms=,num_seasons_actual=,num_years_sim=,num_seasons_sim=,num_obs=,n
um_reps=,outdata=,pmiss=);

    %sim_data(label=&label,apc=&apc,sim_parms=&sim_parms,num_seasons_actual=&num_seasons_actual,n
um_years_sim=&num_years_sim,num_seasons_sim=&num_seasons_sim,num_obs=&num_obs,num_reps=&num_reps,
outdata=sim_data,pmiss=&pmiss);

    %kendall(indata=sim_data,nobs=&num_obs,ny=&num_years_sim,ns=&num_seasons_sim,dv=y,outdata=tes
t_results,pmiss=&pmiss);

    %autocorr(indata=sim_data,medslope=test_results,num_years=&num_years_sim,num_seasons=&num_sea
sons_sim,num_obs=&num_obs,outdata=autocorr);

    data power;
        merge test_results autocorr;
        keep   power_autocorr power_without power_with power_comb;
        retain rej_autocorr rej_without rej_with rej_comb;

        if (_n_=1) then do;
            rej_autocorr=0;
            rej_without =0;
            rej_with    =0;
            rej_comb    =0;
        end;

        rej_autocorr=rej_autocorr+autocorr;
        if (               pwithout1 le 0.05) then rej_without=rej_without+1;
        if (               pwith1    le 0.05) then rej_with   =rej_with   +1;
        if (autocorr=0 and pwithout1 le 0.05) then rej_comb   =rej_comb   +1;
        if (autocorr=1 and pwith1    le 0.05) then rej_comb   =rej_comb   +1;

        if (_n_=&num_reps) then do;
            power_autocorr=rej_autocorr/&num_reps;
            power_without =rej_without /&num_reps;
            power_with    =rej_with    /&num_reps;
            power_comb    =rej_comb    /&num_reps;

            ** Since "with" analysis is not working for missing data, **;
            ** make "with" results missing if there is missing data   **;
            if (&pmiss gt 0) then do;
                power_with=.;
                power_comb=.;
            end;

            output;
        end;
    run;

    ** TEST THE MANN-KENDALL SLOPE ESTIMATOR FOR NORMALITY **;

    proc univariate data=test_results(keep=slope) noprint;
        var slope;
        output out=osl_slope_norm probn=osl_slope_norm;
    run;

    ** CALCULATE THE STANDARD ERROR OF THE MANN-KENDALL SLOPE ESTIMATOR **;

    proc means data=test_results(keep=slope) noprint;
        var slope;
        output out=stderr(drop=_type_ _freq_) std=stderr;
    run;

    ** MERGE ALL SIMULATION SUMMARY RESULTS TOGETHER **;
    ** CALCULATE THE POWER BASED ON THE STDERR VALUE **;

    data &outdata;
        merge &sim_parms power osl_slope_norm stderr;
        length label $25;

        label="&label";
        pmiss=&pmiss;

        b=log(1+&apc);
```

```
        apc=&apc;
        num_seasons_actual=&num_seasons_actual;
        num_years_sim=&num_years_sim;
        num_seasons_sim=&num_seasons_sim;
        num_reps=&num_reps;

        power_stderr=(1-probnorm(probit(0.975)-b/stderr))+probnorm(-probit(0.975)-b/stderr);
        detectable_b=stderr*(probit(0.975)+probit(0.8));
        detectable_apc=exp(detectable_b)-1;
    run;

%mend sim_kendall;

*************************************************************************************;
** CREATE A MACRO TO PERFORM MODELING AND SIMULATION FOR ONE TIME SERIES DATA SET **;
*************************************************************************************;

%macro
run_one_sim(label=,indata=,parameter=,condition=,num_seasons_actual=,num_years_sim=,num_seasons_s
im=,num_obs_sim=,num_reps=,sim_results=,pmiss=);

    ** FIT MIXED MODEL TO GET SIMULATION PARAMETERS **;

    %mixed_model(label=&label,indata=&indata,parameter=&parameter,condition=&condition,num_season
s=&num_seasons_actual,outparms=sim_parms,outdata=mixedout);

    ** SIMULATE THE MANN-KENDALL TREND TEST FOR AN ANNUAL PERCENTAGE CHANGE OF 0 **;

    %sim_kendall(label=&label,apc=0,sim_parms=sim_parms,num_seasons_actual=&num_seasons_actual,nu
m_years_sim=&num_years_sim,num_seasons_sim=&num_seasons_sim,num_obs=&num_obs_sim,num_reps=&num_re
ps,outdata=&sim_results,pmiss=&pmiss);

%mend run_one_sim;

*****************************************************;
** CREATE A MACRO TO ACCUMULATE SIMULATION RESULTS **;
*****************************************************;

%macro accumulate;

    proc datasets nolist;
        append base=all_sims data=sim_results force;
    run;
    quit;

%mend accumulate;

******************;
** MAIN PROGRAM **;
******************;

** SPECIFY LIBRARY WHERE WATER QUALITY TIME SERIES DATA IS STORED **;

libname wqdata "path to folder containing water quality data";

** ROUTE LOG AND OUTPUT TO FILES **;
/*
proc printto log="c:\saslog" print="c:\sasoutput" new; run;
*/

** DELETE ACCUMULATING DATA SET **;

proc datasets library=work nolist;
    delete all_sims;
run;
quit;

** RUN SIMULATIONS **;

%run_one_sim(label=KREA 01  TPO4,indata=wqdata.testdata,parameter=logvalue,condition=project_code
eq "KREA" and station_id eq "KREA 01 " and test_number=25
,num_seasons_actual=24,num_years_sim=5,num_seasons_sim=24,num_obs_sim=120,num_reps=1000,sim_resul
ts=sim_results,pmiss=0.22) %accumulate

%run_one_sim(label=KREA 04  TPO4,indata=wqdata.testdata,parameter=logvalue,condition=project_code
eq "KREA" and station_id eq "KREA 04 " and test_number=25
,num_seasons_actual=12,num_years_sim=5,num_seasons_sim=24,num_obs_sim=120,num_reps=1000,sim_resul
ts=sim_results,pmiss=0.49) %accumulate

%run_one_sim(label=KREA 04  TPO4,indata=wqdata.testdata,parameter=logvalue,condition=project_code
eq "KREA" and station_id eq "KREA 04 " and test_number=25
```

```
,num_seasons_actual=12,num_years_sim=5,num_seasons_sim=24,num_obs_sim=120,num_reps=1000,sim_resul
ts=sim_results,pmiss=0   ) %accumulate

** SAVE SIMULATION RESULTS AS A PERMANENT DATA SET **;

data wqdata.trend_power_results;
    set all_sims;
run;

** PRINT SIMULATION RESULTS **;

proc print data=wqdata.trend_power_results;
    var label num_seasons_actual sigma1sq rho sigma2sq osl_resid_norm
        num_years_sim num_seasons_sim num_reps
        apc power_autocorr power_without power_comb power_with osl_slope_norm stderr power_stderr
detectable_apc pmiss;
    title "Simulation Results";
run;

** IF PREVIOUSLY RE-ROUTED TO FILES, ROUTE LOG AND OUTPUT BACK TO WINDOWS **;
/*
proc printto; run;
*/
```

APPENDIX B to Attachment 1

DETAILED DOCUMENTATION FOR *TREND_POWER.SAS*

The *trend_power.sas* program is a SAS program that is comprised of seven SAS macros followed by a short “main program”. Most users will utilize the program according to the following procedural steps.

A. Create a SAS data set containing the time series data that forms the basis for the power analyses to be performed

   1. Data should be stored in a one time series data point per observation format

   2. The data set should include variables that contain the natural log transformed values of water quality parameters of interest

   3. The data set should contain *season* variable taking on consecutive integer values starting at 1 and defining seasons within each water year for which a fixed seasonal effect is to be included in the mixed model that is fitted to the time series data

B. Specify one or more calls to the *run_one_sim* macro in the “main program” section

   1. Each *run_one_sim* macro call should be followed by an *accumulate* macro call in order to achieve accumulation of the power analysis results from multiple macro calls in a single SAS data set

   2. Results for the basic power analysis procedure may be obtained by executing *run_one_sim* macro calls with the APC argument set to zero

   3. If power estimates are desired for the three variations of the Mann-Kendall trend analysis procedure, these may be obtained by executing *run_one_sim* macro calls with the APC argument set to values other than zero

C. Examine the printed output or perform additional analyses on the permanent SAS data set (*wqdata.trend_power_results*) that contains the power analysis results

Detailed documentation is provided below for each of the seven SAS macros and the main program section included in the *trend_power.sas* program.

1. *mixed_model* macro

A. Functionality – Fits a mixed model to a set of time series data and produces a data set containing the parameters associated with the fitted mixed model

B. Arguments
- label = character string to identify output
- indata = data set contining time series data of interest; must contain a variable named *season* that takes on values 1, 2, … , *num_seasons*
- parameter = name of variable containing natural log transformed time series data of interest
- condition = subsetting condition that selects data of interest from entire data set
- num_seasons = number of seasons represented in the time series data
- outparms = data set to contain the fitted model parameters
- outdata = data set to contain the output data set produced by the MIXED procedure

C. Details
- First attempts to fit a mixed model containing a local nugget; if that model fit does not converge, a model without the local nugget is fitted
- The residuals from the model fit are tested for normality and the observed significance level of the test is saved in the *osl_resid_norm* variable of the *outparms* data set

2. *sim_data* macro

A. Functionality – Simulates a set of time series data based on an input set of mixed model parameters

B. Arguments
- label = character string to identify output
- apc = annual proportion change value to be assumed when simulating data
- sim_parms = data set containing the mixed model parameters used to simulate data
- num_seasons_actual = number of seasons employed in the mixed model that produced the *sim_parms* data set
- num_years_sim = length of time (in years) over which data will be simulated starting at the beginning of 2006
- num_seasons_sim = number of seasons per year for which data is to be simulated
- num_obs = *num_years_sim* times *num_seasons_sim*
- num_reps = number of replicate water quality time series data sets to be simulated
- outdata = data set to contain the simulated data
- pmiss = user-specified proportion of the time that a water quality measurement cannot be obtained

3. *kendall* macro

A. Functionality – Performs the Mann-Kendall trend test on a set of time series data using the kendall3.exe executable Fortran program

B. Arguments
- indata = data set containing the time series data on which the test is to be performed
- nobs = number of observation in the time series

- ny = number of years represented in the time series
- ns = number of seasons per year represented in the time series
- dv = name of dependent variable
- outdata = data set to contain the results of the Mann-Kendall procedure
- pmiss = user-specified proportion of the time that a water quality measurement cannot be obtained (used to suppress some output if pmiss>0)

C. Details – Output data set contains 4 variables
- tau = test statistic value
- pwithout = observed significance level of 2-sided test of the hypothesis that the time series slope is zero, assuming no autocorrelation
- pwith = observed significance level of 2-sided test of the hypothesis that the time series slope is zero, adjusting for autocorrelation
- slope = estimate of the time series slope

4. *autocorr* macro

A. Functionality – Performs a test for serial autocorrelation.

B. Arguments
- indata = name of data set containing simulated data
- medslope = name of data set containing median slope estimators
- num_years = length of time (in years) over which data was simulated
- num_seasons = number of seasons per year for which data was simulated
- num_obs = *num_years* times *num_seasons*
- outdata = name of output data set

C. Details – Rejects the null hypothesis of no serial autocorrlation if and only if the lag1 and lag2 correlation are both statistically significantly positive at a one-sided significance level of 0.05.

5. *sim_kendall* macro

A. Functionality – Based on an input set of mixed model parameters, simulates many sets of time series data and performs the Mann-Kendall trend test for each simulated time series

B. Arguments
- label = character string to identify output
- apc = annual proportion change value to be assumed when simulating data
- sim_parms = data set containing the mixed model parameters used to simulate data
- num_seasons_actual = number of seasons employed in the mixed model that produced the *sim_parms* data set
- num_years_sim = length of time (in years) over which data will be simulated starting at the beginning of 2006
- num_seasons_sim = number of seasons per year for which data is to be simulated
- num_obs = *num_years_sim* times *num_seasons_sim*
- num_reps = number of replicate water quality time series data sets to be simulated
- outdata = data set to contain the Mann-Kendall trend test results for each replicate
- pmiss = user-specified proportion of the time that a water quality measurement cannot be obtained

C. Details - The *outdata* data set contains the *label* variable and the four variables produced by the Mann_Kendall trend procedure (tau, pwithout, pwith, slope)

6. *run_one_sim* macro

A. Functionality – Uses previously defined macros to do the following
- Fit a mixed model to a time series of interest
- Produce a single set of simulated data based on the exact fitted mixed model parameters that may be used to illustrate and check simulation process
- Repeatedly perform the Mann-Kendall procedure for simulate data based on the fitted mixed model parameters and a slope value of zero to examine behavior of the Mann-Kendall procedure under the null hypothesis of a zero slope
- Repeatedly perform the Mann-Kendall procedure for simulate data based on the fitted mixed model parameters and a selected non-zero slope value to examine behavior of the Mann-Kendall procedure under the alternative hypothesis of a non-zero slope

B. Arguments
- label = character string to identify output
- indata = data set containing time series data of interest; must contain a variable named *season* that takes on values 1, 2, … , *num_seasons_actual*
- parameter = name of variable containing time series data of interest
- condition = subsetting condition that selects data of interest from entire data set
- num_seasons_actual = number of seasons represented in the input time series data
- num_years_sim = the number of years over which data is to be simulated starting at the beginning of 2006
- num_seasons_sim = the number of seasons per year for which data is to be simulated
- num_obs_sim = *num_years_sim* times *num_seasons_sim*
- num_reps = number of simulation replications to be performed
- sim_results = name of data set to contain the simulation results

- pmiss = user-specified proportion of the time that a water quality measurement cannot be obtained

C. Details
- Minimum detectable annual proportion change value is determined so that a test for trend with a 2-sided significance level of 5% should reject the null hypothesis 80% of the time

7. *accumulate* macro

A. Functionality – Appends the simulation results data set produced by the *run_one_sim* macro to a permanent data set; individual data sets within are identified by the *label* variable

8. Main Program – The main program performs the following steps:

A. Specify the library where the permanent results data set is to be stored

B. Route log and output to hard disk files (only necessary if many calls to the *run_one_sim* macro are executed)

C. Delete the temporary data set into which simulation results are accumulated

D. Using the *run_one_sim* and *accumulate* macros, run simulations for a series of cases accumulating the simulation results in a temporary SAS data set

E. Save a permanent version of the accumulating temporary SAS data set

F. Print the simulation results

- label = Character string to identify output
- num_seasons_actual = number of seasons represented in the time series data and used in the mixed model
- sigma1sq = estimated mixed model parameter
- rho = estimated mixed model parameter
- sigma2sq = estimated mixed model parameter
- osl_resid_norm = Observed significance level of a test of normality for the residuals from the fixed portion of the fitted mixed model
- num_years_sim = Number of years for which data is simulated
- num_seasons_sim = Number of seasons per year for which data is simulated
- num_reps = Number of water quality time series data sets that are simulated
- apc = Annual proportion change value assumed when simulating data

- power_autocorr = Power of the test for autocorrelation

- power_without = Power of the Mann-Kendall test for trend that applies no correction for serial autocorrelation; if *apc*=0, this value is an estimate of the true significance level of this test procedure

- power_comb = Power of the Mann-Kendall test for trend that applies a correction for serial autocorrelation only if there is statistically significant autocorrelation; if *apc*=0, this value is an estimate of the true significance level of this test procedure

- power_with = Power of the Mann-Kendall test for trend that always applies a correction for serial autocorrelation; if *apc*=0, this value is an estimate of the true significance level of this test procedure

- osl_slope_norm = Observed significance level of a test of normality for the median slope estimator that accompanies the Mann_Kendall procedure

- stderr = Calculated standard deviation of the median slope estimates that accompany the Mann-Kendall procedure

- power_stderr = Power of the test for trend based on the median slope estimator that accompanies the Mann_Kendall procedure; if APC=0, this value will always be 0.05

- detactable_apc = Minimum annual proportion change that is detectable with 80% power using the two-sided 5% test for trend based on the median slope estimator that accompanies the Mann_Kendall procedure

- pmiss = User-specified proportion of the time that a water quality measurement cannot be obtained

G. Re-route log and output to SAS windows

APPENDIX C to Attachment 1

EXAMPLE DATA SET

PROJECT_CODE	STATION_ID	TEST_NUMBER	VALUE	DATE	DATE_MONTHS	LOGVALUE	SEASON
KREA	KREA 01	25	0.035	1992.07	23904.87	-3.3524	2
KREA	KREA 01	25	0.082	1992.11	23905.33	-2.5010	3
KREA	KREA 01	25	0.04	1992.20	23906.34	-3.2189	5
KREA	KREA 01	25	0.034	1992.21	23906.48	-3.3814	5
KREA	KREA 01	25	0.024	1992.32	23907.89	-3.7297	8
KREA	KREA 01	25	0.136	1992.48	23909.75	-1.9951	12
KREA	KREA 01	25	0.088	1992.52	23910.18	-2.4304	13
KREA	KREA 01	25	0.077	1992.56	23910.67	-2.5639	14
KREA	KREA 01	25	0.061	1992.59	23911.10	-2.7969	15
KREA	KREA 01	25	0.144	1992.63	23911.59	-1.9379	16
KREA	KREA 01	25	0.514	1992.67	23912.02	-0.6655	17
KREA	KREA 01	25	0.283	1992.71	23912.48	-1.2623	17
KREA	KREA 01	25	0.117	1992.74	23912.90	-2.1456	18
KREA	KREA 01	25	0.032	1992.78	23913.39	-3.4420	19
KREA	KREA 01	25	0.071	1992.82	23913.82	-2.6451	20
KREA	KREA 01	25	0.054	1992.86	23914.28	-2.9188	21
KREA	KREA 01	25	0.079	1992.90	23914.77	-2.5383	22
KREA	KREA 01	25	0.022	1992.92	23915.10	-3.8167	23
KREA	KREA 01	25	0.027	1992.97	23915.66	-3.6119	24
KREA	KREA 01	25	0.034	1993.05	23916.61	-3.3814	2
KREA	KREA 01	25	0.067	1993.09	23917.10	-2.7031	3
KREA	KREA 01	25	0.004	1993.13	23917.56	-5.5215	4
KREA	KREA 01	25	0.15	1993.17	23917.99	-1.8971	4
KREA	KREA 01	25	0.07	1993.20	23918.42	-2.6593	5
KREA	KREA 01	25	0.323	1993.24	23918.88	-1.1301	6
KREA	KREA 01	25	0.169	1993.28	23919.34	-1.7779	7
KREA	KREA 01	25	0.209	1993.32	23919.80	-1.5654	8
KREA	KREA 01	25	0.116	1993.43	23921.18	-2.1542	11
KREA	KREA 01	25	0.094	1993.51	23922.13	-2.3645	13
KREA	KREA 01	25	0.164	1993.55	23922.56	-1.8079	14
KREA	KREA 01	25	0.086	1993.58	23923.02	-2.4534	15
KREA	KREA 01	25	0.071	1993.62	23923.48	-2.6451	15
KREA	KREA 01	25	0.067	1993.66	23923.94	-2.7031	16
KREA	KREA 01	25	0.134	1993.70	23924.40	-2.0099	17
KREA	KREA 01	25	0.083	1993.74	23924.86	-2.4889	18
KREA	KREA 01	25	0.07	1993.78	23925.42	-2.6593	19
KREA	KREA 01	25	0.059	1993.82	23925.78	-2.8302	20
KREA	KREA 01	25	0.043	1993.86	23926.31	-3.1466	21
KREA	KREA 01	25	0.011	1993.89	23926.70	-4.5099	22
KREA	KREA 01	25	0.026	1993.93	23927.19	-3.6497	23

PROJECT_CODE	STATION_ID	TEST_NUMBER	VALUE	DATE	DATE_MONTHS	LOGVALUE	SEASON
KREA	KREA 01	25	0.088	1994.16	23929.92	-2.4304	4
KREA	KREA 01	25	0.081	1994.20	23930.38	-2.5133	5
KREA	KREA 01	25	0.061	1994.31	23931.76	-2.7969	8
KREA	KREA 01	25	0.036	1994.35	23932.22	-3.3242	9
KREA	KREA 01	25	0.177	1994.47	23933.64	-1.7316	12
KREA	KREA 01	25	0.191	1994.51	23934.13	-1.6555	13
KREA	KREA 01	25	0.077	1994.58	23934.99	-2.5639	14
KREA	KREA 01	25	0.085	1994.66	23935.91	-2.4651	16
KREA	KREA 01	25	0.104	1994.74	23936.83	-2.2634	18
KREA	KREA 01	25	0.403	1994.85	23938.21	-0.9088	21
KREA	KREA 01	25	0.068	1994.93	23939.16	-2.6882	23
KREA	KREA 01	25	0.053	1994.97	23939.69	-2.9375	24
KREA	KREA 01	25	0.036	1995.01	23940.15	-3.3242	1
KREA	KREA 01	25	0.104	1995.05	23940.61	-2.2634	2
KREA	KREA 01	25	0.186	1995.20	23942.42	-1.6820	5
KREA	KREA 01	25	0.128	1995.28	23943.34	-2.0557	7
KREA	KREA 01	25	0.145	1995.32	23943.80	-1.9310	8
KREA	KREA 01	25	0.091	1995.39	23944.72	-2.3969	10
KREA	KREA 01	25	0.255	1995.43	23945.15	-1.3665	11
KREA	KREA 01	25	0.105	1995.47	23945.61	-2.2538	12
KREA	KREA 01	25	0.082	1995.51	23946.13	-2.5010	13
KREA	KREA 01	25	0.072	1995.54	23946.53	-2.6311	14
KREA	KREA 01	25	0.145	1995.59	23947.05	-1.9310	15
KREA	KREA 01	25	0.083	1995.62	23947.45	-2.4889	15
KREA	KREA 01	25	0.118	1995.66	23947.91	-2.1371	16
KREA	KREA 01	25	0.107	1995.69	23948.33	-2.2349	17
KREA	KREA 01	25	0.056	1995.74	23948.83	-2.8824	18
KREA	KREA 01	25	0.071	1995.78	23949.32	-2.6451	19
KREA	KREA 01	25	0.132	1995.81	23949.75	-2.0250	20
KREA	KREA 01	25	0.049	1995.85	23950.21	-3.0159	21
KREA	KREA 01	25	0.057	1995.89	23950.67	-2.8647	22
KREA	KREA 01	25	0.08	1995.92	23951.10	-2.5257	23
KREA	KREA 01	25	0.069	1995.96	23951.56	-2.6736	24
KREA	KREA 01	25	0.051	1996.01	23952.11	-2.9759	1
KREA	KREA 01	25	0.062	1996.05	23952.57	-2.7806	2
KREA	KREA 01	25	0.058	1996.08	23952.97	-2.8473	2
KREA	KREA 01	25	0.081	1996.12	23953.39	-2.5133	3
KREA	KREA 01	25	0.072	1996.16	23953.89	-2.6311	4
KREA	KREA 01	25	0.064	1996.19	23954.31	-2.7489	5
KREA	KREA 01	25	0.04	1996.23	23954.80	-3.2189	6
KREA	KREA 01	25	0.046	1996.27	23955.23	-3.0791	7
KREA	KREA 01	25	0.038	1996.31	23955.69	-3.2702	8
KREA	KREA 01	25	0.03	1996.35	23956.25	-3.5066	9
KREA	KREA 01	25	0.031	1996.38	23956.61	-3.4738	10
KREA	KREA 01	25	0.156	1996.42	23957.07	-1.8579	11
KREA	KREA 01	25	0.279	1996.47	23957.59	-1.2765	12

PROJECT_CODE	STATION_ID	TEST_NUMBER	VALUE	DATE	DATE_MONTHS	LOGVALUE	SEASON
KREA	KREA 01	25	0.173	1996.50	23958.02	-1.7545	13
KREA	KREA 01	25	0.102	1996.55	23958.54	-2.2828	14
KREA	KREA 01	25	0.073	1996.58	23959.00	-2.6173	15
KREA	KREA 01	25	0.037	1996.62	23959.43	-3.2968	15
KREA	KREA 01	25	0.029	1996.65	23959.85	-3.5405	16
KREA	KREA 01	25	0.049	1996.70	23960.34	-3.0159	17
KREA	KREA 01	25	0.022	1996.73	23960.77	-3.8167	18
KREA	KREA 01	25	0.027	1996.77	23961.20	-3.6119	19
KREA	KREA 01	25	0.043	1996.80	23961.66	-3.1466	20
KREA	KREA 01	25	0.025	1996.84	23962.11	-3.6889	21
KREA	KREA 01	25	0.021	1996.89	23962.64	-3.8632	22
KREA	KREA 01	25	0.02	1996.92	23963.03	-3.9120	23
KREA	KREA 01	25	0.022	1996.97	23963.59	-3.8167	24
KREA	KREA 01	25	0.026	1997.00	23963.95	-3.6497	24
KREA	KREA 01	25	0.021	1997.04	23964.51	-3.8632	2
KREA	KREA 01	25	0.019	1997.15	23965.79	-3.9633	4
KREA	KREA 01	25	0.106	1997.46	23969.57	-2.2443	12
KREA	KREA 01	25	0.106	1997.50	23970.03	-2.2443	13
KREA	KREA 01	25	0.123	1997.54	23970.46	-2.0956	13
KREA	KREA 01	25	0.051	1997.57	23970.89	-2.9759	14
KREA	KREA 01	25	0.25	1997.62	23971.41	-1.3863	15
KREA	KREA 01	25	0.102	1997.65	23971.81	-2.2828	16
KREA	KREA 01	25	0.103	1997.69	23972.33	-2.2730	17
KREA	KREA 01	25	0.167	1997.73	23972.73	-1.7898	18
KREA	KREA 01	25	0.121	1997.76	23973.16	-2.1120	19
KREA	KREA 01	25	0.061	1997.81	23973.68	-2.7969	20
KREA	KREA 01	25	0.045	1997.84	23974.08	-3.1011	21
KREA	KREA 01	25	0.032	1997.88	23974.60	-3.4420	22
KREA	KREA 01	25	0.058	1997.92	23975.00	-2.8473	22
KREA	KREA 01	25	0.251	1997.96	23975.49	-1.3823	23
KREA	KREA 01	25	0.163	1997.99	23975.92	-1.8140	24
KREA	KREA 01	25	0.122	1998.03	23976.41	-2.1037	1
KREA	KREA 01	25	0.132	1998.07	23976.84	-2.0250	2
KREA	KREA 01	25	0.174	1998.11	23977.30	-1.7487	3
KREA	KREA 01	25	0.225	1998.15	23977.79	-1.4917	4
KREA	KREA 01	25	0.219	1998.18	23978.22	-1.5187	5
KREA	KREA 01	25	0.309	1998.23	23978.71	-1.1744	6
KREA	KREA 01	25	0.11	1998.26	23979.14	-2.2073	7
KREA	KREA 01	25	0.191	1998.30	23979.63	-1.6555	8
KREA	KREA 01	25	0.3	1998.34	23980.06	-1.2040	9
KREA	KREA 01	25	0.197	1998.38	23980.52	-1.6246	10
KREA	KREA 01	25	0.082	1998.42	23980.98	-2.5010	10
KREA	KREA 01	25	0.056	1998.46	23981.47	-2.8824	11
KREA	KREA 01	25	0.084	1998.49	23981.93	-2.4769	12
KREA	KREA 01	25	0.089	1998.53	23982.36	-2.4191	13
KREA	KREA 01	25	0.421	1998.57	23982.82	-0.8651	14

PROJECT_CODE	STATION_ID	TEST_NUMBER	VALUE	DATE	DATE_MONTHS	LOGVALUE	SEASON
KREA	KREA 01	25	0.449	1998.61	23983.32	-0.8007	15
KREA	KREA 01	25	0.199	1998.65	23983.74	-1.6145	16
KREA	KREA 01	25	0.123	1998.69	23984.27	-2.0956	17
KREA	KREA 01	25	0.257	1998.72	23984.70	-1.3587	18
KREA	KREA 01	25	0.185	1998.76	23985.16	-1.6874	19
KREA	KREA 01	25	0.143	1998.80	23985.58	-1.9449	20
KREA	KREA 01	25	0.103	1998.84	23986.08	-2.2730	21
KREA	KREA 01	25	0.3	1998.88	23986.50	-1.2040	22
KREA	KREA 01	25	0.177	1998.92	23987.00	-1.7316	22
KREA	KREA 01	25	0.07	1998.95	23987.46	-2.6593	23
KREA	KREA 01	25	0.309	1998.99	23987.92	-1.1744	24
KREA	KREA 01	25	0.069	1999.03	23988.35	-2.6736	1
KREA	KREA 01	25	0.093	1999.07	23988.81	-2.3752	2
KREA	KREA 01	25	0.091	1999.11	23989.27	-2.3969	3
KREA	KREA 01	25	0.057	1999.14	23989.73	-2.8647	4
KREA	KREA 01	25	0.052	1999.18	23990.19	-2.9565	5
KREA	KREA 01	25	0.052	1999.22	23990.65	-2.9565	6
KREA	KREA 01	25	0.033	1999.34	23992.06	-3.4112	9
KREA	KREA 01	25	0.041	1999.38	23992.52	-3.1942	10
KREA	KREA 01	25	0.146	1999.49	23993.90	-1.9241	12
KREA	KREA 01	25	0.561	1999.53	23994.36	-0.5780	13
KREA	KREA 01	25	0.505	1999.57	23994.79	-0.6832	14
KREA	KREA 01	25	0.332	1999.60	23995.25	-1.1026	15
KREA	KREA 01	25	0.179	1999.64	23995.71	-1.7204	16
KREA	KREA 01	25	0.178	1999.68	23996.20	-1.7260	17
KREA	KREA 01	25	0.348	1999.72	23996.66	-1.0556	18
KREA	KREA 01	25	0.228	1999.76	23997.09	-1.4784	19
KREA	KREA 01	25	0.401	1999.80	23997.55	-0.9138	20
KREA	KREA 01	25	0.122	1999.83	23998.01	-2.1037	21
KREA	KREA 01	25	0.065	1999.87	23998.47	-2.7334	21
KREA	KREA 01	25	0.065	1999.91	23998.96	-2.7334	22
KREA	KREA 01	25	0.487	1999.95	23999.39	-0.7195	23
KREA	KREA 01	25	0.103	1999.99	23999.85	-2.2730	24
KREA	KREA 01	25	0.085	2000.03	24000.31	-2.4651	1
KREA	KREA 01	25	0.077	2000.11	24001.26	-2.5639	3
KREA	KREA 01	25	0.081	2000.18	24002.18	-2.5133	5
KREA	KREA 01	25	0.078	2000.22	24002.61	-2.5510	6
KREA	KREA 01	25	0.059	2000.26	24003.13	-2.8302	7
KREA	KREA 01	25	0.089	2000.30	24003.56	-2.4191	8
KREA	KREA 01	25	0.51	2000.41	24004.97	-0.6733	10
KREA	KREA 01	25	0.195	2000.45	24005.36	-1.6348	11
KREA	KREA 01	25	0.323	2000.49	24005.85	-1.1301	12
KREA	KREA 01	25	0.123	2000.52	24006.28	-2.0956	13
KREA	KREA 01	25	0.743	2000.56	24006.77	-0.2971	14
KREA	KREA 01	25	1.183	2000.60	24007.20	0.1681	15
KREA	KREA 01	25	1.259	2000.64	24007.69	0.2303	16

PROJECT_CODE	STATION_ID	TEST_NUMBER	VALUE	DATE	DATE_MONTHS	LOGVALUE	SEASON
KREA	KREA 01	25	0.941	2000.68	24008.21	-0.0608	17
KREA	KREA 01	25	0.686	2000.72	24008.61	-0.3769	18
KREA	KREA 01	25	0.745	2000.76	24009.07	-0.2944	19
KREA	KREA 01	25	0.245	2000.79	24009.52	-1.4065	20
KREA	KREA 01	25	0.158	2000.83	24009.98	-1.8452	20
KREA	KREA 01	25	0.091	2000.87	24010.44	-2.3969	21
KREA	KREA 01	25	0.102	2000.91	24010.87	-2.2828	22
KREA	KREA 01	25	0.056	2001.18	24014.15	-2.8824	5
KREA	KREA 01	25	0.131	2001.29	24015.47	-2.0326	7
KREA	KREA 01	25	0.062	2001.37	24016.39	-2.7806	9
KREA	KREA 01	25	0.348	2001.44	24017.31	-1.0556	11
KREA	KREA 01	25	0.503	2001.60	24019.15	-0.6872	15
KREA	KREA 01	25	0.231	2001.63	24019.61	-1.4653	16
KREA	KREA 01	25	0.21	2001.68	24020.14	-1.5606	17
KREA	KREA 01	25	0.179	2001.71	24020.53	-1.7204	18
KREA	KREA 01	25	0.093	2001.76	24021.09	-2.3752	19
KREA	KREA 01	25	0.176	2001.79	24021.45	-1.7373	19
KREA	KREA 01	25	0.324	2001.83	24021.91	-1.1270	20
KREA	KREA 01	25	0.459	2001.87	24022.44	-0.7787	21
KREA	KREA 01	25	0.195	2001.90	24022.83	-1.6348	22
KREA	KREA 01	25	0.286	2001.94	24023.29	-1.2518	23
KREA	KREA 01	25	0.084	2002.02	24024.21	-2.4769	1
KREA	KREA 01	25	0.125	2002.06	24024.77	-2.0794	2
KREA	KREA 01	25	0.069	2002.10	24025.20	-2.6736	3
KREA	KREA 01	25	0.073	2002.17	24026.05	-2.6173	5
KREA	KREA 01	25	0.066	2002.21	24026.52	-2.7181	6
KREA	KREA 01	25	0.06	2002.29	24027.44	-2.8134	7
KREA	KREA 01	25	0.116	2002.48	24029.74	-2.1542	12
KREA	KREA 01	25	0.662	2002.52	24030.20	-0.4125	13
KREA	KREA 01	25	0.318	2002.55	24030.66	-1.1457	14
KREA	KREA 01	25	0.263	2002.59	24031.12	-1.3356	15
KREA	KREA 01	25	0.126	2002.63	24031.58	-2.0715	16
KREA	KREA 01	25	0.233	2002.68	24032.14	-1.4567	17
KREA	KREA 01	25	0.184	2002.71	24032.53	-1.6928	18
KREA	KREA 01	25	0.134	2002.78	24033.42	-2.0099	19
KREA	KREA 01	25	0.14	2002.82	24033.88	-1.9661	20
KREA	KREA 01	25	0.095	2002.90	24034.80	-2.3539	22
KREA	KREA 01	25	0.226	2002.94	24035.33	-1.4872	23
KREA	KREA 01	25	0.246	2002.98	24035.72	-1.4024	24
KREA	KREA 01	25	0.301	2003.02	24036.18	-1.2006	1
KREA	KREA 01	25	0.129	2003.06	24036.67	-2.0479	2
KREA	KREA 01	25	0.113	2003.09	24037.10	-2.1804	3
KREA	KREA 01	25	0.276	2003.13	24037.59	-1.2874	4
KREA	KREA 01	25	0.166	2003.21	24038.48	-1.7958	5
KREA	KREA 01	25	0.108	2003.25	24038.94	-2.2256	6
KREA	KREA 01	25	1.522	2003.44	24041.31	0.4200	11

PROJECT_CODE	STATION_ID	TEST_NUMBER	VALUE	DATE	DATE_MONTHS	LOGVALUE	SEASON
KREA	KREA 01	25	1.094	2003.49	24041.93	0.0898	12
KREA	KREA 01	25	1.07	2003.53	24042.39	0.0677	13
KREA	KREA 01	25	0.444	2003.57	24042.85	-0.8119	14
KREA	KREA 01	25	0.561	2003.61	24043.32	-0.5780	15
KREA	KREA 01	25	0.672	2003.65	24043.78	-0.3975	16
KREA	KREA 01	25	0.446	2003.69	24044.24	-0.8074	17
KREA	KREA 01	25	0.367	2003.73	24044.73	-1.0024	18
KREA	KREA 01	25	0.689	2003.75	24044.99	-0.3725	18
KREA	KREA 01	25	0.445	2003.78	24045.39	-0.8097	19
KREA	KREA 01	25	0.602	2003.82	24045.85	-0.5075	20
KREA	KREA 01	25	0.32	2003.86	24046.31	-1.1394	21
KREA	KREA 01	25	0.457	2003.90	24046.80	-0.7831	22
KREA	KREA 01	25	0.2	2003.94	24047.26	-1.6094	23
KREA	KREA 01	25	0.21	2003.98	24047.72	-1.5606	24
KREA	KREA 01	25	0.176	2004.06	24048.70	-1.7373	2
KREA	KREA 01	25	0.081	2004.09	24049.10	-2.5133	3
KREA	KREA 01	25	0.197	2004.13	24049.59	-1.6246	4
KREA	KREA 01	25	0.126	2004.29	24051.46	-2.0715	7
KREA	KREA 01	25	0.101	2004.32	24051.82	-2.2926	8
KREA	KREA 01	25	0.604	2004.46	24053.52	-0.5042	12
KREA	KREA 01	25	0.709	2004.49	24053.89	-0.3439	12
KREA	KREA 01	25	0.779	2004.61	24055.30	-0.2497	15
KREA	KREA 01	25	0.601	2004.65	24055.75	-0.5092	16
KREA	KREA 01	25	0.384	2004.72	24056.64	-0.9571	18
KREA	KREA 01	25	0.23	2004.76	24057.13	-1.4697	19
KREA	KREA 01	25	0.173	2004.80	24057.62	-1.7545	20
KREA	KREA 01	25	0.206	2004.84	24058.05	-1.5799	21
KREA	KREA 01	25	0.113	2004.88	24058.51	-2.1804	22
KREA	KREA 04	25	0.065	1992.07	23904.87	-2.7334	1
KREA	KREA 04	25	0.121	1992.15	23905.79	-2.1120	2
KREA	KREA 04	25	0.157	1992.20	23906.34	-1.8515	3
KREA	KREA 04	25	0.177	1992.25	23906.97	-1.7316	3
KREA	KREA 04	25	0.287	1992.29	23907.43	-1.2483	4
KREA	KREA 04	25	0.407	1992.32	23907.85	-0.8989	4
KREA	KREA 04	25	0.597	1992.52	23910.18	-0.5158	7
KREA	KREA 04	25	0.306	1992.55	23910.64	-1.1842	7
KREA	KREA 04	25	0.198	1992.59	23911.07	-1.6195	8
KREA	KREA 04	25	0.114	1992.63	23911.52	-2.1716	8
KREA	KREA 04	25	0.111	1992.67	23911.98	-2.1982	8
KREA	KREA 04	25	0.157	1992.71	23912.51	-1.8515	9
KREA	KREA 04	25	0.172	1992.74	23912.93	-1.7603	9
KREA	KREA 04	25	0.101	1992.79	23913.43	-2.2926	10
KREA	KREA 04	25	0.098	1992.82	23913.85	-2.3228	10
KREA	KREA 04	25	0.152	1992.86	23914.31	-1.8839	11
KREA	KREA 04	25	0.201	1992.90	23914.77	-1.6045	11
KREA	KREA 04	25	0.057	1992.94	23915.23	-2.8647	12

PROJECT_CODE	STATION_ID	TEST_NUMBER	VALUE	DATE	DATE_MONTHS	LOGVALUE	SEASON
KREA	KREA 04	25	0.09	1992.98	23915.72	-2.4079	12
KREA	KREA 04	25	0.08	1993.01	23916.15	-2.5257	1
KREA	KREA 04	25	0.116	1993.05	23916.64	-2.1542	1
KREA	KREA 04	25	0.095	1993.09	23917.13	-2.3539	2
KREA	KREA 04	25	0.124	1993.13	23917.59	-2.0875	2
KREA	KREA 04	25	0.104	1993.17	23917.99	-2.2634	2
KREA	KREA 04	25	0.173	1993.20	23918.45	-1.7545	3
KREA	KREA 04	25	0.198	1993.24	23918.91	-1.6195	3
KREA	KREA 04	25	0.099	1993.28	23919.37	-2.3126	4
KREA	KREA 04	25	0.696	1993.32	23919.83	-0.3624	4
KREA	KREA 04	25	0.138	1993.36	23920.29	-1.9805	5
KREA	KREA 04	25	0.256	1993.40	23920.75	-1.3626	5
KREA	KREA 04	25	0.364	1993.47	23921.67	-1.0106	6
KREA	KREA 04	25	0.302	1993.51	23922.16	-1.1973	7
KREA	KREA 04	25	0.189	1993.66	23923.97	-1.6660	8
KREA	KREA 04	25	0.177	1993.70	23924.43	-1.7316	9
KREA	KREA 04	25	0.179	1993.74	23924.89	-1.7204	9
KREA	KREA 04	25	1.191	1993.78	23925.32	0.1748	10
KREA	KREA 04	25	0.12	1993.82	23925.81	-2.1203	10
KREA	KREA 04	25	0.091	1993.89	23926.73	-2.3969	11
KREA	KREA 04	25	0.048	1993.93	23927.16	-3.0366	12
KREA	KREA 04	25	0.03	1993.97	23927.62	-3.5066	12
KREA	KREA 04	25	0.059	1994.01	23928.08	-2.8302	1
KREA	KREA 04	25	0.072	1994.05	23928.58	-2.6311	1
KREA	KREA 04	25	0.053	1994.09	23929.04	-2.9375	2
KREA	KREA 04	25	0.394	1994.13	23929.53	-0.9314	2
KREA	KREA 04	25	0.092	1994.16	23929.96	-2.3860	2
KREA	KREA 04	25	0.1	1994.20	23930.42	-2.3026	3
KREA	KREA 04	25	0.288	1994.24	23930.88	-1.2448	3
KREA	KREA 04	25	0.31	1994.32	23931.80	-1.1712	4
KREA	KREA 04	25	0.249	1994.35	23932.26	-1.3903	5
KREA	KREA 04	25	0.439	1994.47	23933.61	-0.8233	6
KREA	KREA 04	25	0.116	1994.51	23934.10	-2.1542	7
KREA	KREA 04	25	0.14	1994.55	23934.56	-1.9661	7
KREA	KREA 04	25	0.095	1994.63	23935.55	-2.3539	8
KREA	KREA 04	25	0.951	1994.67	23936.04	-0.0502	9
KREA	KREA 04	25	0.107	1994.70	23936.40	-2.2349	9
KREA	KREA 04	25	0.145	1994.74	23936.89	-1.9310	9
KREA	KREA 04	25	0.125	1994.78	23937.32	-2.0794	10
KREA	KREA 04	25	0.087	1994.81	23937.75	-2.4418	10
KREA	KREA 04	25	0.118	1994.85	23938.24	-2.1371	11
KREA	KREA 04	25	0.992	1994.89	23938.70	-0.0080	11
KREA	KREA 04	25	0.263	1996.39	23956.64	-1.3356	5
KREA	KREA 04	25	0.16	1996.43	23957.13	-1.8326	6
KREA	KREA 04	25	0.141	1996.50	23957.98	-1.9590	6
KREA	KREA 04	25	0.119	1996.54	23958.44	-2.1286	7

PROJECT_CODE	STATION_ID	TEST_NUMBER	VALUE	DATE	DATE_MONTHS	LOGVALUE	SEASON
KREA	KREA 04	25	0.21	1996.58	23959.00	-1.5606	8
KREA	KREA 04	25	0.091	1996.62	23959.46	-2.3969	8
KREA	KREA 04	25	0.07	1996.65	23959.82	-2.6593	8
KREA	KREA 04	25	0.115	1996.69	23960.28	-2.1628	9
KREA	KREA 04	25	0.221	1996.73	23960.74	-1.5096	9
KREA	KREA 04	25	0.115	1996.77	23961.26	-2.1628	10
KREA	KREA 04	25	0.068	1996.80	23961.66	-2.6882	10
KREA	KREA 04	25	0.064	1996.88	23962.57	-2.7489	11
KREA	KREA 04	25	0.082	1997.11	23965.33	-2.5010	2
KREA	KREA 04	25	0.169	1997.15	23965.79	-1.7779	2
KREA	KREA 04	25	0.383	1997.20	23966.35	-0.9597	3
KREA	KREA 04	25	0.297	1997.23	23966.71	-1.2140	3
KREA	KREA 04	25	0.257	1997.38	23968.55	-1.3587	5
KREA	KREA 04	25	0.354	1997.46	23969.54	-1.0385	6
KREA	KREA 04	25	0.167	1997.52	23970.20	-1.7898	7
KREA	KREA 04	25	0.195	1997.55	23970.62	-1.6348	7
KREA	KREA 04	25	0.097	1997.60	23971.18	-2.3330	8
KREA	KREA 04	25	0.209	1997.63	23971.55	-1.5654	8
KREA	KREA 04	25	0.13	1997.68	23972.10	-2.0402	9
KREA	KREA 04	25	0.212	1997.71	23972.47	-1.5512	9
KREA	KREA 04	25	0.18	1997.75	23972.99	-1.7148	9
KREA	KREA 04	25	0.097	1997.78	23973.39	-2.3330	10
KREA	KREA 04	25	0.171	1997.82	23973.85	-1.7661	10
KREA	KREA 04	25	0.078	1997.86	23974.31	-2.5510	11
KREA	KREA 04	25	0.09	1997.90	23974.80	-2.4079	11
KREA	KREA 04	25	0.04	1997.94	23975.23	-3.2189	12
KREA	KREA 04	25	0.119	1997.98	23975.72	-2.1286	12
KREA	KREA 04	25	0.075	1998.01	23976.15	-2.5903	1
KREA	KREA 04	25	0.084	1998.06	23976.67	-2.4769	1
KREA	KREA 04	25	0.216	1998.09	23977.07	-1.5325	2
KREA	KREA 04	25	0.094	1998.13	23977.56	-2.3645	2
KREA	KREA 04	25	0.109	1998.17	23977.99	-2.2164	2
KREA	KREA 04	25	0.092	1998.20	23978.45	-2.3860	3
KREA	KREA 04	25	0.123	1998.25	23978.94	-2.0956	3
KREA	KREA 04	25	0.13	1998.28	23979.37	-2.0402	4
KREA	KREA 04	25	0.099	1998.32	23979.83	-2.3126	4
KREA	KREA 04	25	0.134	1998.36	23980.36	-2.0099	5
KREA	KREA 04	25	0.187	1998.40	23980.82	-1.6766	5
KREA	KREA 04	25	0.303	1998.44	23981.28	-1.1940	6
KREA	KREA 04	25	0.323	1998.59	23983.05	-1.1301	8
KREA	KREA 04	25	0.141	1998.63	23983.55	-1.9590	8
KREA	KREA 04	25	0.221	1998.66	23983.97	-1.5096	8
KREA	KREA 04	25	0.173	1998.71	23984.47	-1.7545	9
KREA	KREA 04	25	0.107	1998.74	23984.93	-2.2349	9
KREA	KREA 04	25	0.319	1998.78	23985.39	-1.1426	10
KREA	KREA 04	25	0.091	1998.83	23985.91	-2.3969	10

PROJECT_CODE	STATION_ID	TEST_NUMBER	VALUE	DATE	DATE_MONTHS	LOGVALUE	SEASON
KREA	KREA 04	25	0.502	1998.86	23986.27	-0.6892	11
KREA	KREA 04	25	0.186	1998.89	23986.73	-1.6820	11
KREA	KREA 04	25	0.051	1998.97	23987.65	-2.9759	12
KREA	KREA 04	25	0.056	1999.01	23988.12	-2.8824	1
KREA	KREA 04	25	0.09	1999.05	23988.61	-2.4079	1
KREA	KREA 04	25	0.199	1999.47	23993.64	-1.6145	6
KREA	KREA 04	25	0.302	1999.51	23994.13	-1.1973	7
KREA	KREA 04	25	0.164	1999.55	23994.56	-1.8079	7
KREA	KREA 04	25	0.228	1999.58	23995.02	-1.4784	8
KREA	KREA 04	25	0.295	1999.62	23995.48	-1.2208	8
KREA	KREA 04	25	0.101	1999.67	23996.01	-2.2926	9
KREA	KREA 04	25	0.092	1999.71	23996.50	-2.3860	9
KREA	KREA 04	25	0.185	1999.75	23996.96	-1.6874	9
KREA	KREA 04	25	0.102	1999.78	23997.32	-2.2828	10
KREA	KREA 04	25	0.076	1999.82	23997.78	-2.5770	10
KREA	KREA 04	25	0.063	1999.85	23998.24	-2.7646	11
KREA	KREA 04	25	0.102	1999.89	23998.70	-2.2828	11
KREA	KREA 04	25	0.098	1999.93	23999.16	-2.3228	12
KREA	KREA 04	25	0.097	1999.97	23999.62	-2.3330	12
KREA	KREA 04	25	0.046	2000.09	24001.03	-3.0791	2
KREA	KREA 04	25	0.044	2000.12	24001.46	-3.1236	2
KREA	KREA 04	25	0.323	2000.74	24008.87	-1.1301	9
KREA	KREA 04	25	0.81	2001.60	24019.22	-0.2107	8
KREA	KREA 04	25	0.322	2001.64	24019.64	-1.1332	8
KREA	KREA 04	25	0.196	2001.71	24020.56	-1.6296	9
KREA	KREA 04	25	0.141	2001.75	24021.06	-1.9590	10
KREA	KREA 04	25	0.231	2001.80	24021.55	-1.4653	10
KREA	KREA 04	25	0.142	2001.83	24021.95	-1.9519	10
KREA	KREA 04	25	0.135	2001.87	24022.47	-2.0025	11
KREA	KREA 04	25	0.19	2001.91	24022.90	-1.6607	11
KREA	KREA 04	25	0.131	2001.95	24023.36	-2.0326	12
KREA	KREA 04	25	0.058	2002.02	24024.25	-2.8473	1
KREA	KREA 04	25	0.127	2002.06	24024.74	-2.0636	1
KREA	KREA 04	25	0.125	2002.09	24025.13	-2.0794	2
KREA	KREA 04	25	0.177	2002.17	24026.05	-1.7316	3
KREA	KREA 04	25	0.292	2002.48	24029.77	-1.2310	6
KREA	KREA 04	25	0.201	2002.52	24030.30	-1.6045	7
KREA	KREA 04	25	0.272	2002.55	24030.66	-1.3020	7
KREA	KREA 04	25	0.171	2002.59	24031.12	-1.7661	8
KREA	KREA 04	25	0.126	2002.63	24031.61	-2.0715	8
KREA	KREA 04	25	0.143	2002.68	24032.10	-1.9449	9
KREA	KREA 04	25	0.183	2002.71	24032.53	-1.6983	9
KREA	KREA 04	25	0.17	2002.75	24032.96	-1.7720	9
KREA	KREA 04	25	0.309	2002.86	24034.37	-1.1744	11
KREA	KREA 04	25	0.079	2002.90	24034.83	-2.5383	11
KREA	KREA 04	25	0.076	2002.94	24035.26	-2.5770	12

PROJECT_CODE	STATION_ID	TEST_NUMBER	VALUE	DATE	DATE_MONTHS	LOGVALUE	SEASON
KREA	KREA 04	25	0.067	2002.98	24035.82	-2.7031	12
KREA	KREA 04	25	0.061	2003.02	24036.18	-2.7969	1
KREA	KREA 04	25	0.048	2003.06	24036.67	-3.0366	1
KREA	KREA 04	25	0.08	2003.09	24037.10	-2.5257	2
KREA	KREA 04	25	0.069	2003.14	24037.63	-2.6736	2
KREA	KREA 04	25	0.185	2003.21	24038.48	-1.6874	3
KREA	KREA 04	25	0.169	2003.25	24038.94	-1.7779	3
KREA	KREA 04	25	0.181	2003.48	24041.80	-1.7093	6
KREA	KREA 04	25	0.232	2003.51	24042.16	-1.4610	7
KREA	KREA 04	25	0.133	2003.56	24042.72	-2.0174	7
KREA	KREA 04	25	0.118	2003.59	24043.08	-2.1371	8
KREA	KREA 04	25	0.116	2003.63	24043.55	-2.1542	8
KREA	KREA 04	25	0.162	2003.67	24044.04	-1.8202	9
KREA	KREA 04	25	0.141	2003.71	24044.53	-1.9590	9
KREA	KREA 04	25	0.107	2003.75	24044.99	-2.2349	9
KREA	KREA 04	25	0.137	2003.78	24045.39	-1.9878	10
KREA	KREA 04	25	0.105	2003.82	24045.85	-2.2538	10
KREA	KREA 04	25	0.039	2003.98	24047.72	-3.2442	12
KREA	KREA 04	25	0.049	2004.09	24049.10	-3.0159	2
KREA	KREA 04	25	0.12	2004.18	24050.21	-2.1203	3
KREA	KREA 04	25	0.1	2004.29	24051.46	-2.3026	4
KREA	KREA 04	25	0.143	2004.57	24054.84	-1.9449	7
KREA	KREA 04	25	0.32	2004.61	24055.30	-1.1394	8
KREA	KREA 04	25	0.221	2004.65	24055.75	-1.5096	8
KREA	KREA 04	25	0.2	2004.72	24056.64	-1.6094	9
KREA	KREA 04	25	0.175	2004.76	24057.13	-1.7430	10
KREA	KREA 04	25	0.153	2004.80	24057.62	-1.8773	10
KREA	KREA 04	25	0.107	2004.84	24058.05	-2.2349	11
KREA	KREA 06A	25	0.139	1992.49	23909.92	-1.9733	6
KREA	KREA 06A	25	0.519	1992.54	23910.44	-0.6559	7
KREA	KREA 06A	25	0.339	1992.57	23910.84	-1.0818	7
KREA	KREA 06A	25	0.253	1992.61	23911.30	-1.3744	8
KREA	KREA 06A	25	0.226	1992.65	23911.79	-1.4872	8
KREA	KREA 06A	25	0.363	1992.69	23912.28	-1.0134	9
KREA	KREA 06A	25	0.32	1992.72	23912.67	-1.1394	9
KREA	KREA 06A	25	0.266	1992.76	23913.16	-1.3243	10
KREA	KREA 06A	25	0.137	1992.80	23913.62	-1.9878	10
KREA	KREA 06A	25	0.606	1992.84	23914.08	-0.5009	11
KREA	KREA 06A	25	0.099	1992.88	23914.54	-2.3126	11
KREA	KREA 06A	25	0.084	1992.92	23915.03	-2.4769	12
KREA	KREA 06A	25	0.096	1992.95	23915.46	-2.3434	12
KREA	KREA 06A	25	0.171	1992.99	23915.92	-1.7661	12
KREA	KREA 06A	25	0.16	1993.03	23916.38	-1.8326	1
KREA	KREA 06A	25	0.09	1993.07	23916.81	-2.4079	1
KREA	KREA 06A	25	0.185	1993.11	23917.30	-1.6874	2
KREA	KREA 06A	25	0.209	1993.15	23917.82	-1.5654	2

PROJECT_CODE	STATION_ID	TEST_NUMBER	VALUE	DATE	DATE_MONTHS	LOGVALUE	SEASON
KREA	KREA 06A	25	0.177	1993.19	23918.25	-1.7316	3
KREA	KREA 06A	25	0.418	1993.22	23918.68	-0.8723	3
KREA	KREA 06A	25	0.371	1993.26	23919.14	-0.9916	4
KREA	KREA 06A	25	0.381	1993.30	23919.60	-0.9650	4
KREA	KREA 06A	25	0.202	1993.38	23920.52	-1.5995	5
KREA	KREA 06A	25	0.27	1993.42	23921.01	-1.3093	6
KREA	KREA 06A	25	0.23	1993.46	23921.51	-1.4697	6
KREA	KREA 06A	25	0.186	1993.49	23921.90	-1.6820	6
KREA	KREA 06A	25	0.212	1993.53	23922.36	-1.5512	7
KREA	KREA 06A	25	0.276	1993.57	23922.82	-1.2874	7
KREA	KREA 06A	25	0.228	1993.61	23923.28	-1.4784	8
KREA	KREA 06A	25	0.226	1993.65	23923.74	-1.4872	8
KREA	KREA 06A	25	0.255	1993.69	23924.24	-1.3665	9
KREA	KREA 06A	25	0.168	1993.76	23925.09	-1.7838	10
KREA	KREA 06A	25	0.247	1993.80	23925.55	-1.3984	10
KREA	KREA 06A	25	0.087	1993.84	23926.04	-2.4418	11
KREA	KREA 06A	25	0.137	1993.87	23926.47	-1.9878	11
KREA	KREA 06A	25	0.096	1993.91	23926.96	-2.3434	11
KREA	KREA 06A	25	0.104	1993.95	23927.39	-2.2634	12
KREA	KREA 06A	25	0.115	1993.99	23927.85	-2.1628	12
KREA	KREA 06A	25	0.162	1994.03	23928.31	-1.8202	1
KREA	KREA 06A	25	0.144	1994.06	23928.77	-1.9379	1
KREA	KREA 06A	25	0.14	1994.10	23929.23	-1.9661	2
KREA	KREA 06A	25	0.226	1994.14	23929.73	-1.4872	2
KREA	KREA 06A	25	0.171	1994.18	23930.15	-1.7661	3
KREA	KREA 06A	25	0.295	1994.22	23930.61	-1.2208	3
KREA	KREA 06A	25	0.205	1994.26	23931.07	-1.5847	4
KREA	KREA 06A	25	0.271	1994.29	23931.53	-1.3056	4
KREA	KREA 06A	25	0.238	1994.34	23932.03	-1.4355	5
KREA	KREA 06A	25	0.182	1994.45	23933.41	-1.7037	6
KREA	KREA 06A	25	0.377	1994.49	23933.87	-0.9755	6
KREA	KREA 06A	25	0.232	1994.53	23934.33	-1.4610	7
KREA	KREA 06A	25	0.191	1994.57	23934.79	-1.6555	7
KREA	KREA 06A	25	0.17	1994.60	23935.25	-1.7720	8
KREA	KREA 06A	25	0.187	1994.64	23935.71	-1.6766	8
KREA	KREA 06A	25	0.261	1994.69	23936.27	-1.3432	9
KREA	KREA 06A	25	0.603	1994.76	23937.12	-0.5058	10
KREA	KREA 06A	25	0.5	1994.80	23937.58	-0.6931	10
KREA	KREA 06A	25	0.183	1994.84	23938.04	-1.6983	11
KREA	KREA 06A	25	0.213	1994.87	23938.47	-1.5465	11
KREA	KREA 06A	25	0.223	1994.91	23938.96	-1.5006	11
KREA	KREA 06A	25	0.169	1994.95	23939.42	-1.7779	12
KREA	KREA 06A	25	0.178	1994.99	23939.85	-1.7260	12
KREA	KREA 06A	25	0.129	1995.03	23940.35	-2.0479	1
KREA	KREA 06A	25	0.119	1995.07	23940.81	-2.1286	1
KREA	KREA 06A	25	0.192	1995.15	23941.76	-1.6503	2

PROJECT_CODE	STATION_ID	TEST_NUMBER	VALUE	DATE	DATE_MONTHS	LOGVALUE	SEASON
KREA	KREA 06A	25	0.204	1995.18	23942.12	-1.5896	3
KREA	KREA 06A	25	0.343	1995.22	23942.58	-1.0700	3
KREA	KREA 06A	25	0.219	1995.25	23943.04	-1.5187	4
KREA	KREA 06A	25	0.201	1995.30	23943.63	-1.6045	4
KREA	KREA 06A	25	0.143	1995.33	23943.96	-1.9449	4
KREA	KREA 06A	25	0.275	1995.37	23944.42	-1.2910	5
KREA	KREA 06A	25	0.109	1995.45	23945.38	-2.2164	6
KREA	KREA 06A	25	0.113	1995.48	23945.80	-2.1804	6
KREA	KREA 06A	25	0.112	1995.52	23946.26	-2.1893	7
KREA	KREA 06A	25	0.109	1995.56	23946.72	-2.2164	7
KREA	KREA 06A	25	0.173	1995.60	23947.18	-1.7545	8
KREA	KREA 06A	25	0.131	1995.64	23947.64	-2.0326	8
KREA	KREA 06A	25	0.235	1995.68	23948.17	-1.4482	9
KREA	KREA 06A	25	0.259	1995.72	23948.66	-1.3509	9
KREA	KREA 06A	25	0.295	1995.76	23949.12	-1.2208	10
KREA	KREA 06A	25	0.97	1995.79	23949.48	-0.0305	10
KREA	KREA 06A	25	0.151	1995.84	23950.04	-1.8905	11
KREA	KREA 06A	25	0.101	1995.87	23950.41	-2.2926	11
KREA	KREA 06A	25	0.114	1995.91	23950.96	-2.1716	11
KREA	KREA 06A	25	0.066	1995.94	23951.33	-2.7181	12
KREA	KREA 06A	25	0.055	1995.98	23951.82	-2.9004	12
KREA	KREA 06A	25	0.05	1996.02	23952.28	-2.9957	1
KREA	KREA 06A	25	0.092	1996.06	23952.74	-2.3860	1
KREA	KREA 06A	25	0.072	1996.10	23953.23	-2.6311	2
KREA	KREA 06A	25	0.123	1996.14	23953.72	-2.0956	2
KREA	KREA 06A	25	0.114	1996.18	23954.15	-2.1716	3
KREA	KREA 06A	25	0.121	1996.22	23954.61	-2.1120	3
KREA	KREA 06A	25	0.206	1996.25	23955.03	-1.5799	4
KREA	KREA 06A	25	0.143	1996.30	23955.56	-1.9449	4
KREA	KREA 06A	25	0.272	1996.33	23955.92	-1.3020	4
KREA	KREA 06A	25	0.138	1996.37	23956.48	-1.9805	5
KREA	KREA 06A	25	0.375	1996.41	23956.93	-0.9808	5
KREA	KREA 06A	25	0.228	1996.45	23957.39	-1.4784	6
KREA	KREA 06A	25	0.204	1996.48	23957.75	-1.5896	6
KREA	KREA 06A	25	0.243	1996.52	23958.21	-1.4147	7
KREA	KREA 06A	25	0.244	1996.56	23958.67	-1.4106	7
KREA	KREA 06A	25	0.18	1996.60	23959.23	-1.7148	8
KREA	KREA 06A	25	0.162	1996.63	23959.59	-1.8202	8
KREA	KREA 06A	25	0.166	1996.68	23960.15	-1.7958	9
KREA	KREA 06A	25	0.144	1996.71	23960.51	-1.9379	9
KREA	KREA 06A	25	0.159	1996.76	23961.07	-1.8389	10
KREA	KREA 06A	25	0.157	1996.79	23961.52	-1.8515	10
KREA	KREA 06A	25	0.161	1996.83	23961.92	-1.8264	10
KREA	KREA 06A	25	0.093	1996.90	23962.80	-2.3752	11
KREA	KREA 06A	25	0.135	1996.94	23963.26	-2.0025	12
KREA	KREA 06A	25	0.076	1996.98	23963.72	-2.5770	12

PROJECT_CODE	STATION_ID	TEST_NUMBER	VALUE	DATE	DATE_MONTHS	LOGVALUE	SEASON
KREA	KREA 06A	25	0.112	1997.02	23964.18	-2.1893	1
KREA	KREA 06A	25	0.054	1997.06	23964.67	-2.9188	1
KREA	KREA 06A	25	0.099	1997.10	23965.20	-2.3126	2
KREA	KREA 06A	25	0.095	1997.13	23965.59	-2.3539	2
KREA	KREA 06A	25	0.19	1997.18	23966.12	-1.6607	3
KREA	KREA 06A	25	0.128	1997.33	23967.96	-2.0557	4
KREA	KREA 06A	25	0.15	1997.36	23968.32	-1.8971	5
KREA	KREA 06A	25	0.14	1997.48	23969.74	-1.9661	6
KREA	KREA 06A	25	0.135	1997.52	23970.20	-2.0025	7
KREA	KREA 06A	25	0.242	1997.55	23970.62	-1.4188	7
KREA	KREA 06A	25	0.145	1997.60	23971.18	-1.9310	8
KREA	KREA 06A	25	0.362	1997.63	23971.55	-1.0161	8
KREA	KREA 06A	25	0.17	1997.68	23972.10	-1.7720	9
KREA	KREA 06A	25	0.27	1997.71	23972.47	-1.3093	9
KREA	KREA 06A	25	0.229	1997.75	23972.99	-1.4740	9
KREA	KREA 06A	25	0.138	1997.78	23973.39	-1.9805	10
KREA	KREA 06A	25	0.124	1997.82	23973.85	-2.0875	10
KREA	KREA 06A	25	0.087	1997.86	23974.31	-2.4418	11
KREA	KREA 06A	25	0.126	1997.90	23974.80	-2.0715	11
KREA	KREA 06A	25	0.085	1997.94	23975.23	-2.4651	12
KREA	KREA 06A	25	0.272	1997.98	23975.72	-1.3020	12
KREA	KREA 06A	25	0.158	1998.01	23976.15	-1.8452	1
KREA	KREA 06A	25	0.186	1998.06	23976.67	-1.6820	1
KREA	KREA 06A	25	0.154	1998.09	23977.07	-1.8708	2
KREA	KREA 06A	25	0.277	1998.17	23977.99	-1.2837	2
KREA	KREA 06A	25	0.179	1998.20	23978.45	-1.7204	3
KREA	KREA 06A	25	0.377	1998.25	23978.94	-0.9755	3
KREA	KREA 06A	25	0.22	1998.28	23979.37	-1.5141	4
KREA	KREA 06A	25	0.235	1998.32	23979.83	-1.4482	4
KREA	KREA 06A	25	0.297	1998.36	23980.36	-1.2140	5
KREA	KREA 06A	25	0.332	1998.40	23980.82	-1.1026	5
KREA	KREA 06A	25	0.296	1998.44	23981.28	-1.2174	6
KREA	KREA 06A	25	0.254	1998.48	23981.74	-1.3704	6
KREA	KREA 06A	25	0.582	1998.55	23982.62	-0.5413	7
KREA	KREA 06A	25	0.374	1998.59	23983.05	-0.9835	8
KREA	KREA 06A	25	0.406	1998.63	23983.55	-0.9014	8
KREA	KREA 06A	25	0.378	1998.66	23983.97	-0.9729	8
KREA	KREA 06A	25	0.233	1998.71	23984.47	-1.4567	9
KREA	KREA 06A	25	0.449	1998.78	23985.39	-0.8007	10
KREA	KREA 06A	25	0.194	1998.83	23985.91	-1.6399	10
KREA	KREA 06A	25	0.348	1998.86	23986.27	-1.0556	11
KREA	KREA 06A	25	0.285	1998.89	23986.73	-1.2553	11
KREA	KREA 06A	25	0.189	1998.97	23987.65	-1.6660	12
KREA	KREA 06A	25	0.181	1999.01	23988.12	-1.7093	1
KREA	KREA 06A	25	0.229	1999.05	23988.61	-1.4740	1
KREA	KREA 06A	25	0.23	1999.09	23989.04	-1.4697	2

PROJECT_CODE	STATION_ID	TEST_NUMBER	VALUE	DATE	DATE_MONTHS	LOGVALUE	SEASON
KREA	KREA 06A	25	0.102	1999.13	23989.53	-2.2828	2
KREA	KREA 06A	25	0.12	1999.16	23989.96	-2.1203	2
KREA	KREA 06A	25	0.235	1999.47	23993.64	-1.4482	6
KREA	KREA 06A	25	0.555	1999.51	23994.13	-0.5888	7
KREA	KREA 06A	25	0.385	1999.55	23994.56	-0.9545	7
KREA	KREA 06A	25	0.333	1999.58	23995.02	-1.0996	8
KREA	KREA 06A	25	0.31	1999.62	23995.48	-1.1712	8
KREA	KREA 06A	25	0.341	1999.67	23996.01	-1.0759	9
KREA	KREA 06A	25	0.435	1999.71	23996.50	-0.8324	9
KREA	KREA 06A	25	0.384	1999.75	23996.96	-0.9571	9
KREA	KREA 06A	25	0.793	1999.78	23997.32	-0.2319	10
KREA	KREA 06A	25	0.757	1999.82	23997.78	-0.2784	10
KREA	KREA 06A	25	0.256	1999.85	23998.24	-1.3626	11
KREA	KREA 06A	25	0.292	1999.89	23998.70	-1.2310	11
KREA	KREA 06A	25	0.233	1999.93	23999.16	-1.4567	12
KREA	KREA 06A	25	0.281	1999.97	23999.62	-1.2694	12
KREA	KREA 06A	25	0.189	2000.01	24000.08	-1.6660	1
KREA	KREA 06A	25	0.132	2000.05	24000.61	-2.0250	1
KREA	KREA 06A	25	0.137	2000.09	24001.03	-1.9878	2
KREA	KREA 06A	25	0.172	2000.12	24001.46	-1.7603	2
KREA	KREA 06A	25	0.393	2000.24	24002.87	-0.9339	3
KREA	KREA 06A	25	0.566	2000.58	24006.97	-0.5692	7
KREA	KREA 06A	25	0.429	2000.62	24007.43	-0.8463	8
KREA	KREA 06A	25	0.39	2000.66	24007.95	-0.9416	8
KREA	KREA 06A	25	0.281	2000.74	24008.87	-1.2694	9
KREA	KREA 06A	25	0.287	2000.77	24009.26	-1.2483	10
KREA	KREA 06A	25	0.298	2001.48	24017.77	-1.2107	6
KREA	KREA 06A	25	0.379	2001.52	24018.26	-0.9702	7
KREA	KREA 06A	25	0.571	2001.60	24019.22	-0.5604	8
KREA	KREA 06A	25	0.494	2001.64	24019.64	-0.7052	8
KREA	KREA 06A	25	0.437	2001.68	24020.17	-0.8278	9
KREA	KREA 06A	25	0.272	2001.71	24020.56	-1.3020	9
KREA	KREA 06A	25	0.205	2001.75	24021.06	-1.5847	10
KREA	KREA 06A	25	0.324	2001.80	24021.55	-1.1270	10
KREA	KREA 06A	25	0.343	2001.83	24021.95	-1.0700	10
KREA	KREA 06A	25	0.317	2001.87	24022.47	-1.1489	11
KREA	KREA 06A	25	0.263	2001.91	24022.90	-1.3356	11
KREA	KREA 06A	25	0.224	2001.95	24023.36	-1.4961	12
KREA	KREA 06A	25	0.089	2002.02	24024.25	-2.4191	1
KREA	KREA 06A	25	0.223	2002.06	24024.74	-1.5006	1
KREA	KREA 06A	25	0.223	2002.09	24025.13	-1.5006	2
KREA	KREA 06A	25	0.21	2002.17	24026.05	-1.5606	3
KREA	KREA 06A	25	0.198	2002.48	24029.77	-1.6195	6
KREA	KREA 06A	25	0.5	2002.52	24030.30	-0.6931	7
KREA	KREA 06A	25	0.528	2002.55	24030.66	-0.6387	7
KREA	KREA 06A	25	0.427	2002.59	24031.12	-0.8510	8

PROJECT_CODE	STATION_ID	TEST_NUMBER	VALUE	DATE	DATE_MONTHS	LOGVALUE	SEASON
KREA	KREA 06A	25	0.315	2002.63	24031.61	-1.1552	8
KREA	KREA 06A	25	0.394	2002.68	24032.10	-0.9314	9
KREA	KREA 06A	25	0.331	2002.71	24032.53	-1.1056	9
KREA	KREA 06A	25	0.278	2002.75	24032.96	-1.2801	9
KREA	KREA 06A	25	0.306	2002.79	24033.45	-1.1842	10
KREA	KREA 06A	25	0.374	2002.83	24033.91	-0.9835	10
KREA	KREA 06A	25	0.097	2002.90	24034.83	-2.3330	11
KREA	KREA 06A	25	0.15	2002.94	24035.26	-1.8971	12
KREA	KREA 06A	25	0.169	2002.98	24035.82	-1.7779	12
KREA	KREA 06A	25	0.217	2003.02	24036.18	-1.5279	1
KREA	KREA 06A	25	0.085	2003.06	24036.67	-2.4651	1
KREA	KREA 06A	25	0.089	2003.09	24037.10	-2.4191	2
KREA	KREA 06A	25	0.135	2003.14	24037.63	-2.0025	2
KREA	KREA 06A	25	0.312	2003.21	24038.48	-1.1648	3
KREA	KREA 06A	25	0.18	2003.25	24038.94	-1.7148	3
KREA	KREA 06A	25	0.148	2003.29	24039.47	-1.9105	4
KREA	KREA 06A	25	0.219	2003.36	24040.36	-1.5187	5
KREA	KREA 06A	25	0.256	2003.48	24041.80	-1.3626	6
KREA	KREA 06A	25	0.414	2003.51	24042.16	-0.8819	7
KREA	KREA 06A	25	0.25	2003.56	24042.72	-1.3863	7
KREA	KREA 06A	25	0.368	2003.59	24043.08	-0.9997	8
KREA	KREA 06A	25	0.486	2003.63	24043.55	-0.7215	8
KREA	KREA 06A	25	0.477	2003.67	24044.04	-0.7402	9
KREA	KREA 06A	25	0.418	2003.71	24044.53	-0.8723	9
KREA	KREA 06A	25	0.482	2003.82	24045.85	-0.7298	10
KREA	KREA 06A	25	0.308	2003.90	24046.80	-1.1777	11
KREA	KREA 06A	25	0.16	2004.09	24049.10	-1.8326	2
KREA	KREA 06A	25	0.202	2004.13	24049.59	-1.5995	2
KREA	KREA 06A	25	0.136	2004.29	24051.46	-1.9951	4
KREA	KREA 06A	25	0.176	2004.46	24053.52	-1.7373	6
KREA	KREA 06A	25	0.195	2004.49	24053.89	-1.6348	6
KREA	KREA 06A	25	0.282	2004.53	24054.38	-1.2658	7
KREA	KREA 06A	25	0.208	2004.57	24054.84	-1.5702	7
KREA	KREA 06A	25	0.206	2004.61	24055.30	-1.5799	8
KREA	KREA 06A	25	0.652	2004.65	24055.75	-0.4277	8
KREA	KREA 06A	25	0.287	2004.88	24058.51	-1.2483	11
KREA	KREA 06A	25	0.24	2004.91	24058.93	-1.4271	11

ATTACHMENT 2: PROJECT-SPECIFIC OPTIMIZATION REPORTS

These documents are available as a separate pdf files from the SFWMD.

Projects East of Lake Okeechobee
IRL
SE
WQM

Projects North of Lake Okeechobee
KREA
LKR
TCNS
V

Projects West of Lake Okeechobee
BRM
CCWQ
CESWQ
CR

Projects in Lake Okeechobee
OLIT
Y

Projects in the Everglades Agricultural Area
CAMB
SEMI
ST1W

Projects on the Southwest Coast of Florida
BISC

ATTACHMENT 3: PROGRESS REPORTS AND DOTT MEETING SUMMARIES

These documents are available as a separate pdf file from the SFWMD.